04/34578

(12) INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

(19) World Intellectual Property Organization
International Bureau

(43) International Publication Date
3 February 2005 (03.02.2005)

PCT

(10) International Publication Number
WO 2005/010034 A1

(51) International Patent Classification[7]: C07K 14/165, A61K 39/215, 39/42, 38/16

(21) International Application Number: PCT/US2004/023345

(22) International Filing Date: 20 July 2004 (20.07.2004)

(25) Filing Language: English

(26) Publication Language: English

(30) Priority Data:
60/489,166 21 July 2003 (21.07.2003) US
60/524,642 25 November 2003 (25.11.2003) US

(71) Applicant (*for all designated States except US*): NATIONAL INSTITUTES OF HEALTH [US/US]; 6011 Executive Boulevard, Suite 325, Rockville, Maryland 20852-3804 (US).

(72) Inventors; and
(75) Inventors/Applicants (*for US only*): DIMITROV, Dimiter S. [US/US]; 12309 Braxfield Court, Apartment #16, Rockville, Maryland 20852 (US). XIAO, Xiaodong [US/US]; 2201 Bannerhill Road, Frederick, Maryland 21702 (US).

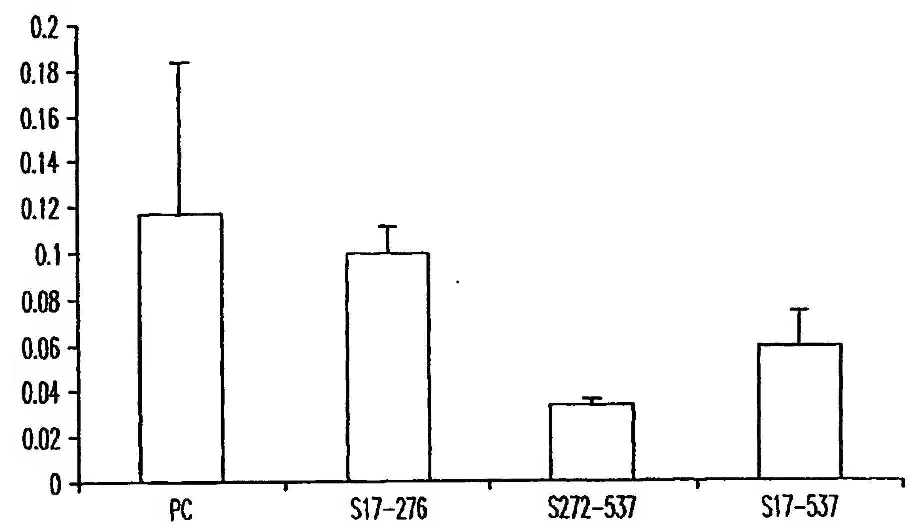
(74) Agents: STEFFEY, Charles, E. et al.; P.O. Box 2938, Minneapolis, MN 55402 (US).

(81) Designated States (*unless otherwise indicated, for every kind of national protection available*): AE, AG, AL, AM, AT, AU, AZ, BA, BB, BG, BR, BW, BY, BZ, CA, CH, CN, CO, CR, CU, CZ, DE, DK, DM, DZ, EC, EE, EG, ES, FI, GB, GD, GE, GH, GM, HR, HU, ID, IL, IN, IS, JP, KE, KG, KP, KR, KZ, LC, LK, LR, LS, LT, LU, LV, MA, MD, MG, MK, MN, MW, MX, MZ, NA, NI, NO, NZ, OM, PG, PH, PL, PT, RO, RU, SC, SD, SE, SG, SK, SL, SY, TJ, TM, TN, TR, TT, TZ, UA, UG, US, UZ, VC, VN, YU, ZA, ZM, ZW.

(84) Designated States (*unless otherwise indicated, for every kind of regional protection available*): ARIPO (BW, GH, GM, KE, LS, MW, MZ, NA, SD, SL, SZ, TZ, UG, ZM, ZW), Eurasian (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM),

[Continued on next page]

(54) Title: SOLUBLE FRAGMENTS OF THE SARS-COV SPIKE GLYCOPROTEIN



(57) Abstract: The invention relates generally to the spike protein from the virus (SARS-CoV) that is etiologically linked to severe acute respiratory syndrome (SARS); polypeptides and peptide fragments of the spike protein, and conservative variants thereof; nucleic acid segments and constructs that encode the spike protein, polypeptides and peptide fragments of the spike protein, and conservative variants thereof, and coupled proteins that include the spike protein or a portion thereof; peptidomimetics; vaccines; methods for vaccination and treatment of severe acute respiratory syndrome; antibodies; aptamers; and kits containing antibodies or aptamers that bind to the spike protein.

WO 2005/010034 A1

European (AT, BE, BG, CH, CY, CZ, DE, DK, EE, ES, FI, FR, GB, GR, HU, IE, IT, LU, MC, NL, PL, PT, RO, SE, SI, SK, TR), OAPI (BF, BJ, CF, CG, CI, CM, GA, GN, GQ, GW, ML, MR, NE, SN, TD, TG).

Published:

— *with international search report*

— *before the expiration of the time limit for amending the claims and to be republished in the event of receipt of amendments*

For two-letter codes and other abbreviations, refer to the "Guidance Notes on Codes and Abbreviations" appearing at the beginning of each regular issue of the PCT Gazette.

Soluble Fragments of the SARS-CoV Spike Glycoprotein

This application claims priority from U.S. Application Ser. No. 60/489,166 filed July 21, 2003 and from U.S. Application Ser. No. 60/524,642 filed November 25, 2003, which are hereby incorporated by reference in their entireties.

Government Funding

The invention described herein was developed with the support of the Department of Health and Human Services. The United States Government has certain rights in the invention.

Field of the Invention

The invention relates generally to a spike polypeptide that is encoded by a coronavirus (herein SARS-CoV), which is etiologically linked to Severe Acute Respiratory Syndrome (SARS). The invention further relates to nucleic acids and polypeptides having amino acid sequences that correspond to fragments of spike protein of SARS-CoV, and conservative variants thereof. The invention also relates to use of these nucleic acids, polypeptides, variants, and fragments to produce antibodies that recognize the spike protein of SARS-CoV, and for the production of vaccines against SARS. Another aspect of the invention relates to spike protein fragments for inhibiting fusion of the SARS-CoV with animal cells.

Background of the Invention

Severe acute respiratory syndrome (SARS) is an infectious atypical pneumonia that has recently been recognized in patients in 32 countries and regions. The atypical pneumonia with unknown etiology was initially observed in Guangdong Province, China. This observation was followed by reports from Hong Kong, Vietnam, Singapore, Canada and Beijing of severe febrile respiratory illness that spread to household members and health care workers. This disease was later designated "severe acute respiratory syndrome (SARS)" by the World Health Organization (WHO). Until May 19, 2003, a cumulative

total of 7,864 SARS cases were reported to WHO from 29 countries. A total of 643 deaths (case-fatality proportion: 8.2 %) were reported.

Researchers around the world have sequenced the genome of SARS causing viruses from different regions of the globe. The viruses have been classified as coronaviruses. Coronaviruses have been grouped into three categories based on cross-reactivity of antibodies backed up by genetic data. Two previously known human viruses fell into different groups than SARS-CoV. The coronavirus that causes SARS does not fit into any of the previously known clusters. Rather, it forms a new group by itself. Phylogenetic analysis of the predicted viral proteins indicates that the virus does not closely resemble any of the three previously known groups of coronaviruses. Most coronaviruses cause either a respiratory or an enteric disease, which is also transmitted by the faecal-oral route.

The incubation period for SARS is usually 2 to 7 days. Infection is characterized by fever, non-productive cough, shortness of breath, and the presence of minimal auscultatory findings with consolidation on chest radiographs. Lymphopenia, leucopenia, thrombocytopenia, and elevated liver enzymes and creatinine kinase may also be present in most cases. Symptoms relating to the gastrointestinal tract were also noticed in SARS patients.

Pathological studies of patients who died of SARS from Guangdong, Hongkong, Beijing and Singapore showed diffuse alveolar damage (DAD) in the lung as the most notable feature. In those individuals with severe disease resulting in death, scattered type II pneumocytes showed marked cytologic changes that include multinucleation, cytomegaly, nucleomegaly, clearing of nuclear chromatin, and prominent nucleoli. Although these changes were severe, they were within the spectrum of epithelial changes seen in other cases of diffuse alveolar damage. Morphologic changes that were identified included bronchial epithelial denudation, loss of cilia, and squamous metaplasia. Other findings included focal intraalveolar hemorrhage, hemophagocytosis, necrotic inflammatory debris in small airways, organizing pneumonia or secondary bacterial pneumonia.

The pathogenesis of this disorder remains to be determined. However, the mechanism of acute lung injury could involve direct damage by the virus to the alveolar wall by targeting either endothelial cells or epithelial cells.

Alternatively, the virus could infect inflammatory cells with the injury mediated through cytokines, interleukins, or tumor necrosis factor-alpha. It is also possible that the tissue damage in SARS is not directly related to viral infection in tissues but is a secondary effect of cytokines or other factors induced by viral infection proximal to but not within the lung tissue.

Pathologic evaluation of the fatal cases showed that hepatocytes underwent fatty degeneration, cloudy swelling, apoptosis and dot necrosis, with Kupffer cell proliferation and portal infiltrates of lymphocytes. There were regional hemorrhages, vascular congestion and lymphocytic infiltration in gastrointestinal walls of the patient.

Due to the ability of SARS-CoV to be spread through an airborne route, SARS-CoV presents a particular threat to the health of large populations of people throughout the world. Accordingly, methods to immunize people before infection, diagnose infection, immunize people during infection, and treat infected persons infected with SARS-CoV are greatly needed.

Summary of the Invention

These and other needs are met by the invention described herein. The invention provides polypeptides; peptide fragments; viral fusion inhibitors; coupled proteins; immunopeptides; immune compositions; peptidomimetics; nucleic acid segments; expression cassettes; nucleic acid constructs; recombinant viruses; viral vaccines; peptide vaccines; microorganism vaccines; DNA vaccines; antibodies; aptamers; pharmaceutical compositions; methods to immunize an animal; a method to treat severe acute respiratory syndrome (SARS); methods to diagnose SARS; and kits.

The invention provides polypeptides having an amino acid sequence corresponding to that of a polypeptide that is etiologically linked to SARS. Preferably the polypeptide is the spike protein from SARS-CoV that can inhibit SARS fusion with animal cells and/or raise immune response in an animal. In some embodiments, the polypeptide is a soluble form of the spike protein from SARS-CoV. In other embodiments, the polypeptide includes amino acids 17-757 of the spike protein from SARS-CoV. In some embodiments, the polypeptide includes amino acids 762-1189 of the spike protein from SARS-CoV. In other embodiments, the polypeptide includes amino acids 17-757 of the

spike protein from SARS-CoV. In some embodiments, the polypeptide includes amino acids 17-276 of the spike protein from SARS-CoV. In other embodiments, the polypeptide includes amino acids 303-537 of the spike protein from SARS-CoV. In some embodiments, the polypeptide includes amino acids 317-517 of the spike protein from SARS-CoV. In other embodiments, the polypeptide includes amino acids 272-537 of the spike protein from SARS-CoV. In some embodiments, the polypeptide includes amino acids 17-537 of the spike protein from SARS-CoV. In other embodiments, the polypeptide includes amino acids 17-1189 (relative to SEQ ID NO: 1) of the spike protein from SARS-CoV. The polypeptides of the invention can inhibit SARS-CoV fusion with animal cells. The nucleic acids and polypeptides of the invention can elicit an immune response when used to inoculate an animal. In some embodiments, the nucleic acids and polypeptides of the invention elicit a cellular immune response when used to inoculate an animal. In other embodiments, the nucleic acids and polypeptides of the invention elicit a humoral immune response when used to inoculate an animal. The animal can be a reptile. In some embodiments, the animal is an avian. In other embodiments, the animal is a mammal. Sometimes, the animal is a human.

The invention provides peptide fragments of the spike protein from SARS-CoV. Preferably the peptide fragments are soluble in aqueous solution. A peptide fragment of the invention may lack one amino acid residue from the amino acid sequence of the full length spike protein from SARS-CoV. In some embodiments, peptide fragments are at least three amino acids in length. In other embodiments, peptide fragments are at least 10 amino acids in length. In some embodiments, peptide fragments are at least 20 amino acids in length. In other embodiments, peptide fragments are at least 30 amino acids in length. In some embodiments, peptide fragments are at least 40 amino acids in length. In other embodiments, peptide fragments are at least 50 amino acids in length. In some embodiments, peptide fragments are at least 60 amino acids in length. The peptide fragments may also be single amino acid unit additions to a fragment of a given length. For example, peptide fragment may be 3, 4, 10, 11, 21, 22, 31, or 32 amino acids in length. The peptide fragments of the invention can inhibit SARS Co-V fusion with animal cells or elicit an immune response when used to inoculate an animal. Examples of peptides that can elicit an immune response

after inoculation of an animal include, for example, the D24 peptide having sequence DVQAPNYTQHTSSMRGC (SEQ ID NO:58) and the P540 peptide having sequence PSSKRFQPQQFGRDC (SEQ ID NO:59). In some embodiments, the peptide fragments of the invention elicit a cellular immune response when used to inoculate an animal. In other embodiments, the peptide fragments of the invention elicit a humoral immune response when used to inoculate an animal. The animal can be a reptile. In some embodiments, the animal is an avian. In other embodiments, the animal is a mammal. In further embodiments, the animal is a human.

The invention provides coupled proteins. The coupled proteins include a carrier protein that is coupled to a second polypeptide. Preferably, the carrier protein is soluble. In some embodiments, the carrier protein increases an immune response to the second polypeptide of the coupled protein when used to inoculate an animal. In other embodiments, the carrier protein elicits a cellular immune response to the second polypeptide of the coupled protein when used to inoculate an animal. In some embodiments, the carrier protein elicits a humoral immune response to the second polypeptide of the coupled protein when used to inoculate an animal. The second polypeptide can be a polypeptide or a peptide fragment of the invention, or a conservative variant thereof. The animal can be a reptile. In some embodiments, the animal is an avian. In other embodiments, the animal is a mammal. In further embodiments, the animal is a human.

The invention provides immunopeptides that include a polypeptide or peptide fragment of the invention, or a conservative variant thereof, that is coupled to an acetyl group, a picryl group, an arsanilic acid, or to a sulfanilic acid. In some embodiments, the immunopeptide is coupled to an acetyl or a picryl group. In other embodiments, immunopeptide is coupled to arsanilic acid or sulfanilic acid. Preferably, the immunopeptide is soluble. Preferably, the immunopeptide elicits an immune response when used to inoculate an animal. In some embodiments, the immunopeptide elicits a humoral immune response when used to inoculate an animal. In other embodiments, the immunopeptide elicits a cellular immune response when used to inoculate an animal. The animal can be a reptile. In some embodiments, the animal is an avian. In other embodiments, the animal is a mammal. In further embodiments, the animal is a human.

The invention provides peptidomimetics that are polypeptides or peptide fragments of the invention, and conservative variants thereof, in which a peptide bond has been replaced with a non-peptide bond. In some embodiments, the peptidomimetic can inhibit SARS Co-V fusion with animal cells. In other embodiments, the peptidomimetic elicits an immune response when used to inoculate an animal. For example, the peptidomimetic can elicit a cellular immune response when used to inoculate an animal. Alternatively, the peptidomimetic elicits a humoral immune response when used to inoculate an animal. The animal can be a reptile. In some embodiments, the animal is an avian. In other embodiments, the animal is a mammal. In further embodiments, the animal is a human.

The invention provides compositions containing an adjuvant and a nucleic acid, polypeptide, a peptide fragment, or a peptidomimetic of the invention. In some embodiments, the composition inhibits SARS-CoV fusion with animal cells. In other embodiments, the composition elicits an immune response when used to inoculate an animal. In some embodiments, the immune composition elicits a cellular immune response when used to inoculate an animal. In other embodiments, the immune composition elicits a humoral immune response when used to inoculate an animal. The animal can be a reptile. In some embodiments, the animal is an avian. In other embodiments, the animal is a mammal. In further embodiments, the animal is a human.

The invention provides nucleic acid segments that encode polypeptides and peptide fragments of the invention, and conservative variants thereof.

The invention provides expression cassettes having a promoter that is operably linked to a nucleic acid segment of the invention. In some embodiments, the promoter is constitutive. In other embodiments, the promoter is inducible.

The invention provides nucleic acid constructs that include a vector and a nucleic acid segment of the invention. The nucleic acid construct can include an expression cassette of the invention. In some embodiments, the vector can be a virus. In other embodiments, the vector is a plasmid. In further embodiments, the vector is an expression vector.

The invention provides a recombinant virus that includes a viral vector and a nucleic acid segment of the invention. In some embodiments, the viral

vector is a herpes virus. In other embodiments, the viral vector is a canarypox virus. In other embodiments, the viral vector is an adenovirus. In further embodiments, the viral vector is a vaccinia virus.

The invention provides a viral vaccine against SARS that includes a viral vector, a nucleic acid segment of the invention, and a pharmaceutical carrier. In some embodiments, the viral vector is a herpes virus. In other embodiments, the viral vector is a canarypox virus. In other embodiments, the viral vector is an adenovirus. In further embodiments, the viral vector is a vaccinia virus. Preferably, the pharmaceutical carrier is formulated for injection. Preferably, the viral vaccine elicits an immune response when used to inoculate an animal. In some embodiments, the viral vaccine elicits a cellular immune response when used to inoculate an animal. In other embodiments, the viral vaccine elicits a humoral immune response when used to inoculate an animal. The animal can be a reptile. In some embodiments, the animal is an avian. In other embodiments, the animal is a mammal. In further embodiments, the animal is a human.

The invention provides a peptide vaccine against SARS that includes a peptidomimetic, polypeptide or a peptide fragment of the invention, or a conservative variant thereof, and a pharmaceutical carrier. Preferably, the pharmaceutical carrier is formulated for injection. Preferably, the peptide vaccine is formulated in unit dosage form. Preferably, the peptide vaccine elicits an immune response when used to inoculate an animal. In some embodiments, the peptide vaccine elicits a cellular immune response when used to inoculate an animal. In other embodiments, the peptide vaccine elicits a humoral immune response when used to inoculate an animal. The animal can be a reptile. In some embodiments, the animal is an avian. In other embodiments, the animal is a mammal. In further embodiments, the animal is a human.

The invention provides a microorganism vaccine against SARS that includes a microorganism that expresses a polypeptide or a peptide fragment of the invention, or a conservative variant thereof, and a pharmaceutical carrier. Preferably, the microorganism is attenuated. In some embodiments, the microorganism is Salmonella. In other embodiments, the microorganism is Listeria. In further embodiments, the microorganism is Listeria monocytogenes. In some embodiments, the pharmaceutical carrier is formulated for injection. In other embodiments, the pharmaceutical carrier is formulated for oral

administration. Preferably, the microorganism vaccine is formulated in unit dosage form. Preferably, the microorganism vaccine elicits an immune response when used to inoculate an animal. In some embodiments, the microorganism vaccine elicits a cellular immune response when used to inoculate an animal. In other embodiments, the microorganism vaccine elicits a humoral immune response when used to inoculate an animal. The animal can be a reptile. In some embodiments, the animal is an avian. In other embodiments, the animal is a mammal. In further embodiments, the animal is a human.

The invention provides a DNA vaccine against SARS that includes a vector into which is inserted a nucleic acid segment of the invention, and a pharmaceutical carrier. The DNA vaccine may include an adjuvant. The DNA vaccine may include a myonecrotic agent. For example, the myonecrotic agent can be bupivicaine. In other embodiments, the myonecrotic agent is cardiotoxin. The vector can, for example, be a virus. In other embodiments, the vector is a bacteriophage. In further embodiments, the vector is a plasmid. The vector containing the insert can be prepared in a eukaryotic cell. However, in some embodiments, the vector containing the insert is prepared in a prokaryotic cell. For example, the vector containing the insert can be prepared in a bacterium. In some embodiments, the pharmaceutical carrier is formulated for mucosal delivery. In other embodiments, the pharmaceutical carrier is formulated for injection. Preferably, the DNA vaccine is formulated in unit dosage form. Preferably, the DNA vaccine elicits an immune response when used to inoculate an animal. In some embodiments, the DNA vaccine elicits a humoral immune response when used to inoculate an animal. In other embodiments, the DNA vaccine elicits a cellular immune response when used to inoculate an animal. The animal can be a reptile. In some embodiments, the animal is an avian. In other embodiments, the animal is a mammal. In further embodiments, the animal is a human.

The invention provides an antibody that binds to a polypeptide or peptide fragment of the invention, or a conservative variant thereof. In some embodiments, the antibody is an antigen-binding antibody fragment. In other embodiments, the antibody is a polyclonal antibody. In further embodiments, the antibody is a single-chain antibody. In other embodiments, the antibody is a monoclonal antibody. In some preferred embodiments, the antibody is a

humanized antibody. The antibody may be coupled to a detectable tag. For example, the detectable tag can be a radiolabel. In some embodiments, the detectable tag is an affinity tag. In other embodiments, the detectable tag is an enzyme. In further embodiments, the detectable tag is a fluorescent protein. In some preferred embodiments, the detectable tag is a fluorescent marker. The antibody may also be coupled to a toxin.

The invention provides aptamers that bind to a polypeptide or peptide fragment of the invention, or a conservative variant thereof. The aptamer may be coupled to a detectable tag. For example, the detectable tag is a radiolabel. In some embodiments, the detectable tag is an affinity tag. In other embodiments, the detectable tag is an enzyme. In further embodiments, the detectable tag is a fluorescent protein. In some preferred embodiments, the detectable tag is a fluorescent marker. The aptamer may also be coupled to a toxin.

The invention provides a pharmaceutical composition or a kit containing an antibody, S polypeptide or aptamer of the invention and a pharmaceutical carrier. Preferably, the pharmaceutical composition is formulated for injection.

Brief Description of the Figures

This patent or application file contains at least one drawing executed in color. Copies of this patent or patent application publication with color drawing(s) will be provided by the Office upon request and payment of the necessary fee.

Fig. 1A illustrates an agarose gel electrophoresis of a DNA construct having an insert that encodes the spike protein of the invention. Lanes from left to right: Lane 1 is a one kb DNA ladder (markers from bottom to top – 0.5, 1.0, 1.6, 2.0, 3.0, 4.0); Lane 2 shows the DNA construct digested with BamHI/XbaI, resulting in the distinctive vector band (upper band) and the DNA fragment that encodes the spike protein (lower band); Lane 3 shows the DNA construct digested with HindIII which produced a smaller band and a larger band as expected due to the presence of a HindIII site in the vector and within the DNA fragment encoding the spike protein.

Fig. 1B provides a schematic diagram of a monomer of the full-length SARS-CoV S glycoprotein showing various soluble polypeptide fragments after

removal of the signal sequence (residues 1–16, SEQ ID NO:60). The soluble fragments are spike protein fragments named "S" followed by numbers corresponding to the spike protein amino acids that constitute the termini of the fragment. Thus, "S756" is a soluble spike protein fragment beginning at amino acid 17 (just after the signal sequence) and ending at amino acid 756. "TM" denotes the transmembrane segment and the arrow indicates a possible cleavage site within amino acid positions 758–761 (sequence RNTR). "RBD" indicates the potential receptor-binding domain that is within amino acid positions 272–537 (SEQ ID NO:57), likely between a residue downstream from position 303 and a residue upstream of position 537 (SEQ ID NO:61).

Fig. 2 illustrates a denaturing polyacrylamide gel electrophoresis (SDS-PAGE) of the expression of a peptide fragment of the spike protein from SARS-CoV in Escherichia coli. The peptide fragment corresponds to amino acids 17-446 of SEQ ID NO: 1. The nucleic acid segment encoding amino acids 17-446 was cloned into a pRSET vector to create pRSET-S(17-446), which was expressed in BL21DE3 cells. Numbers and arrows on the left indicate molecular weight markers in kilodaltons. The lanes contain the following polypeptides: M - molecular weight markers; lanes 1 and 2 - polypeptides of control E. coli containing the pRSET vector without the nucleic acid segment encoding amino acid residues 17-446 of SEQ ID NO: 1 and without isopropylthiogalactoside (IPTG) induction; lane 3 - polypeptides of control E. coli containing the pRSET vector without the nucleic acid segment encoding amino acid residues 17-446 of SEQ ID NO: 1 but with IPTG induction; lane 4 - analysis of E. coli containing the pRSET vector with a nucleic acid segment encoding amino acid residues 17-446 of SEQ ID NO: 1, and with IPTG induction. The arrow on the right side indicates the position of a peptide fragment corresponding to amino acid residues 17-446 of SEQ ID NO: 1 as expressed in E. coli.

Fig. 3 illustrates a slot blot analysis of the expression of the indicated peptide fragments of the spike protein from SARS-CoV in mammalian cells. Nucleic acid segments coding for the peptide fragments were cloned into a pSecTag2B vector to express peptide fragments having the mouse k chain leader sequence at the N-terminus for secretion, and a c-Myc epitope plus a histidine tag at the C-terminus for detection and affinity purification. The nucleic acid constructs were transformed into HEK293 and VeroE6 cells. Expression of the

indicated peptide fragments was examined through use of slot blot analysis with an anti-c-Myc antibody. The numbers on the left and right indicate the amino acid residues included within the detected peptide fragments. The left column represents expression of the peptide fragments in HEK293 cells. The right column represents expression of the peptide fragments in VeroE6 cells. The upper half represents samples obtained from medium in which the cells were grown (secreted proteins), and the lower half represents samples obtained from cell lysate (intracellular portion). PC is a positive control, provided by the manufacturer of the plasmid that contains PSA with a c-Myc tag at the C-terminus. NC is a negative control that contains the full length spike protein from SARS-CoV that lacks a c-Myc epitope or histidine tag.

Fig. 4A illustrates a slot blot analysis of the expression of the indicated peptide fragments from the spike protein from SARS-CoV in human 293 or Monkey VeroE6 cells. Supernatants of 293 and Vero E6 cells transfected with plasmids encoding S fragments (S276, S537, and S756) in the absence or presence of T7 polymerase expressed by recombinant vaccinia virus (VTF7.3) were transferred to nitrocellulose membranes and detected with anti-c-Myc epitope antibody. The numbers on the left and right indicate the amino acid residues included within the detected peptide fragments. PSA PC is a positive control that contains PSA with a c-Myc tag at the C-terminus. pCDNA-S NC is a negative control that contains the full length spike protein from SARS-CoV that lacks a c-Myc epitope or histidine tag. The lanes are as follows: (1) human 293 cells that were not infected with a VTF7.3 vaccinia virus, (2) human 293 cells that were infected with a VTF7.3 vaccinia virus, (3) monkey VeroE6 cells that were not infected with a VTF7.3 vaccinia virus, and (4) monkey VeroE6 cells that were infected with a VTF7.3 vaccinia virus.

Fig. 4B Supernatants from transfected cells as described above for Fig. 4A were incubated with Ni–NTA agarose beads, washed, and subjected to Western blotting with the same anti-c-Myc epitope antibody as in Fig. 4A.

Fig. 4C illustrates detection of S fragments by two rabbit polyclonal antibodies raised against peptides corresponding to sequences starting at residues 24 (D24, middle panel) and 540 (P540, right panel), respectively. The left panel shows for comparison Western blot where S537 and S756 were detected by the anti-c-Myc epitope antibody.

Fig. 5 illustrates that the full-length membrane-associated S protein is expressed on the surface of cells, as shown by flow cytometry using the rabbit polyclonal antibody P540. A nucleic acid encoding the full-length S glycoprotein was used to transfect 293 cells, which were then infected with VTF7.3. Cells were collected and incubated with P540 polyclonal antibody plus anti-rabbit secondary antibody conjugated with FITC, washed, and subjected to flow cytometry analysis. The same plasmid used to express S but without the nucleic acids for S was used to transfect cells in a control experiment denoted as negative control (NC); cells with nucleic acids encoding the full-length S glycoprotein are denoted as S.

Fig. 6A and 6B illustrate that substantially no cleavage of the S glycoprotein occurs naturally. Western blots of supernatants from transfected 293 cells expressing S756, Se, and cell lysate of 293 cells expressing the S glycoproteins using the P540 antibody are shown. Close to background level cleavage of S and Se was observed. Fig. 6A shows a Western blot of samples kept for three days at 4 °C before analysis to monitor the effect of nonspecific protease activity on the cleavage pattern. In contrast, Fig. 6B shows blots with samples used immediately after preparation.

Fig. 7A-C shows that cell fusion is mediated by the S glycoprotein. A pCDNA3-based plasmid without S insert was used as plasmid control, and fusion between S-expressing cells with ACE2-ecto expressing cells was used as negative control. The pCDNA3-ACE2-ecto construct expresses just the ACE2 soluble ecto domain tagged with C9 peptide. Fig.7A illustrates that there was no syncytium formation between 293T cells transfected with pSecTag2B-S and pCDNA3-ACE2-Ecto. In contrast, Fig.7B illustrates syncytium formation between 293T cells transfected with pSecTag2B-S and pCDNA3-ACE2, respectively. Fig. 7C graphically illustrates cell fusion as measured by a reporter gene-based assay. As shown, S glycoprotein expressed in both pCDNA3 and pSecTag2B vectors can be detected in a β-gal reporter gene-based cell–cell fusion assay.

Fig. 8A-C shows that the S glycoprotein receptor-binding domain (RBD) is localized between residues 272 and 537. Fig. 8A illustrates binding of two different S soluble fragments (S537 and S756) to 293 and Vero E6 cells. Fig. 8B illustrates binding of various S fragments to Vero E6 cells. The background

OD_{405} measured for the negative control was subtracted from the OD_{405} values of each S fragment. The resulting OD_{405} for each fragment was then presented as a percentage of the OD_{405} for S537. Fig. 8C illustrates which S polypeptide fragments interact with purified soluble ACE2 as measured by ELISA. In all experiments, the negative control (NC) represents sample processed exactly the same way as the others except that the plasmid used for transfection did not encode any protein. Data shown here represent at least three independent experiments. OD_{405} for all samples is presented as percentages of the OD_{405} for S537.

Fig. 9A-D illustrates that dimerization occurs between the N terminal fragments of the SARS-CoV S glycoprotein as demonstrated by co-immunoprecipitation and cross-linking. All N-terminal fragments except the smallest fragment (S317-517) containing the receptor binding domain were coimmunoprecipiated with S756 by the P540 antibody. The P540 antibody is a rabbit polyclonal antibody that was developed against a peptide containing residues 540-555 of the S glycoprotein and it binds the S756 polypeptide but not the N-terminal fragments.

In Fig. 9A, plasmids encoding N-terminal fragments (denoted by the number of the ending amino acid residue or the number of the starting and ending residue) were used to transfect 293T cells alone (left six lanes) or in combination (right four lanes) with a S756-encoding plasmid. These cells were then infected with the vaccinia virus VTF7.3. After incubation, the culture medium was collected and subjected to Western blot analysis using a mouse anti-c-Myc epitope antibody that recognizes all fragments.

Fig. 9B shows that all N-terminal S fragments, except the smallest fragment (S317-517) that contained the receptor binding domain, were coimmunoprecipiated with S756 by the P540 antibody. The same medium samples used in Fig. 9A were subjected first to immunoprecipitation with the P540 polyclonal antibody that recognizes only S756. These immunoprecipitates were then subjected to Western blot analysis using the anti-c-Myc epitope antibody to confirm that the N-terminal fragments coimmunoprecipitated.

Fig. 9C shows that a new band with a molecular weight corresponding to a dimer forms in the presence and absence of DTT. To rule out the possibility of nonspecific disulfide bond formation that may lead to coimmunoprecipitation,

DTT was included in one of the coimmunoprecipitation experiment. DTT had no effect on either immunoprecipitation or coimmunoprecipitation of secreted S756 (left lanes) or S756+S276 (right lanes). Medium samples containing secreted S756 (left lanes) or S756+S276 (right lanes) fragments were subjected to immunoprecipitation with P540 in the presence or absence of 2 mM DTT.

Fig. 9D illustrates the size of the S polypeptide oligomers. The S537 fragment was cross-linked with BS^3 (Pierce, Rockford, IL) as described in the Examples and a Western blot was prepared after SDS-PAGE separation and the anti-c-Myc antibody was used for detection of the S537 monomer and its oligomers. As shown in the right lane of Fig. 9D, a new band appears when the crosslinking reagent is added. The new band had a molecular weight corresponding to a dimer but not of higher order oligomers.

Fig. 10A illustrates dimerization of the N terminal fragment S537 as detected by size-exclusion chromatography. The elution profiles of S537 and S317-517 are shown with arrows and numbers indicating the position and molecular weight at which standard calibration proteins were eluted.

Fig. 10B provides western blots of fractions collected for S537 and S317-517 by using an anti-c-Myc epitope antibody.

Fig. 11A-B illustrates that the extreme N terminal domain is required for the S glycoprotein mediated cell-cell fusion. Fig. 11A provides a schematic representation of the S glycoprotein deletion mutants and a summary of the data from a cell-cell fusion assay where RBD denotes the approximate position of the receptor binding domain. The presence of signal due to fusion is denoted by a plus (+) and lack of measurable signal above background levels by a minus (-). Only wild type polypeptides with amino acids 17-1255 had fusion activity. Neither of the deletion mutants having amino acids 103-1255 (Del1) or 311-1255 (Del2) had fusion activity. Fig. 11B shows the levels of expression of full length and deletion mutants of the S glycoprotein as measured by Western analysis. Equal amount of cell lysates were loaded for each sample and the rabbit polyclonal antibody P540 was used for detection. Fig. 11C illustrates that the full length S glycoprotein and the Del1 and Del 2 deletion mutants are expressed on the cell surface as measured by flow cytometry. The level of surface expression was low although the negative control where the cells were

transfected with an empty plasmid was clearly distinguishable to the left of the other three curves.

Fig. 12A-B illustrates that dimeric S1 binds more efficiently to the receptor ACE2 than monovalent fragments containing the receptor binding domain. Fig. 12A shows the relative levels of expression of different S fragments as detected by ELISA using 200 μl of culture supernatants from cells transfected with S276, S319-518 and S537 constructs. Anti-His and anti-c-Myc epitope antibodies were used in a sandwich ELISA to detect the levels of secreted tagged S proteins. Fig. 12B shows the level of binding by S fragments to ACE2 as measured by ELISA. The tagged ACE2 was bound to plates by an anti-C9 antibody that had been previously coated on the plates. The supernatants from cell cultures where the cells were transfected with various S proteins were mixed and incubated in ELISA plates either with (hatched bars) or without (open bars) anti-c-Myc antibody. The highest level of expression or binding is assumed to be 100 %. As shown the S537 fragment with both the N-terminal dimerization domain and the receptor binding domain, binds ACE2 more efficiently than does the S319-518 fragment that has only the receptor binding domain.

Fig. 13A-B illustrates that the soluble S ectodomain is trimeric under the conditions of size-exclusion chromatography. In Fig. 13A, purified Se was run on a gel filtration column that was calibrated by using proteins with known molecular weight. BSA in equal amount was included as an internal control. In Fig. 13B, different fractions were collected from the gel filtration column and analyzed by Western blot. Two bands S polypeptide are detected in some fractions that contain Se fragments of the indicated molecular weights, representing the Se fragment alone (lower band) and its aggregates (upper band).

Fig. 14A illustrates that a DNA vaccine of the invention can elicit very high titer anti-SARS-CoV sera in mice. Mice 1A-5A were immunized with DNA encoding the S319-518 fragment that contains the spike protein receptor binding domain (RBD). Mice 1B-5B were immunized with RBD-encoding DNA (the S319-518 fragment) fused to a nucleic acid encoding an Fc fragment. Mice 1C-3C received plasmid only (no S fragment DNA). Anti-sera were collected and tested via ELISA to ascertain the titer of the different isolates. In Fig. 14A, the first number denotes an individual mouse, the letter denotes the

respective immunization group, and the last number denotes the dilution used. Anti-sera were diluted by factors of 50, 250, 1250 and 7250, as shown on the x-axis of the bar graph. These data indicate that immunization with DNA encoding the receptor binding domain of the S protein induces a strong immune response against SARS-CoV.

Fig. 14B illustrates that anti-sera from mice immunized with RBD-encoding DNA can prevent S-mediated cell fusion. Cells (293T) were incubated with anti-sera from mice immunized with DNA encoding a spike protein receptor binding domain polypeptide (S319-518) fragment and then the cell suspension was mixed with cells expressing S protein. Fusion was measured as described in Example 20 (see also, Xiao et al. BBRC 2003). The percentage (where 1=100%) of activity for each fusion reaction is plotted on the y-axis, where the percentage of the fusion without any inhibition was designated as 100%. PC denotes positive control where no serum was added. For mice sera #1 to #2 in each group, serum dilution factors of 10 (designated 0.1), 100 (designated 0.01), and 1000 (designated (0.001) were used. For mice sera #3-#5 in groups A and B, and #3 in the control group, dilution factors of 20 (designated 0.05) and 100 (designated 0.01) were used. These data indicate that immunization with DNA encoding the receptor binding domain of the S protein could prevent SARS-CoV infection.

Fig. 15 illustrates that soluble S glycoprotein fragments inhibit S-mediated cell fusion. 10 ug/ml of various S fragments were incubated with ACE2-expressing cells first for 10 min at room temperature. The ACE2-expressing cells were then mixed with S expressing cells and the fusion assay was carried out as described in the Examples. The Y-axis is the OD_{595} for each sample after the background noise was subtracted. Numbers of each construct represent the starting and ending residues of the respective polypeptide.

Detailed Description of the Invention

SARS represents an important public health concern. Methods to diagnose and treat persons who are infected with SARS-CoV provide the opportunity to either prevent or control further spread of infection by SARS-CoV. These methods are especially important due to the ability of SARS-CoV to infect persons through an airborne route. The present invention provides

nucleic acids that encode segments of the amino acid sequence of the spike protein of SARS-CoV. The present invention also provides polypeptides that correspond in amino acid sequence to segments of the amino acid sequence of the spike protein of SARS-CoV. The invention also provides peptide fragments and conservative variants of the spike protein of SARS-CoV, in addition to coupled proteins and peptidomimetics that have portions which correspond in amino acid sequence to the spike protein.

The spike protein is important because it is present on the outside of intact SARS-CoV. Thus, it presents a target that can be used to inhibit or eliminate an intact virus before the virus has an opportunity to infect a cell.

The nucleic acids and polypeptides of the invention offer advantages over the full length spike protein because the nucleic acids are easy to produce and the polypeptides of the invention are produced in large amounts in soluble form. The polypeptides of the invention offer additional advantages over the native spike protein because they can be made to have increased resistant to degradation when administered to an animal. The polypeptides of the invention can also be formulated to increase their antigenicity to make them more efficient antigens to elicit an immune response when administered to an animal, such as a human.

Accordingly, the invention provides nucleic acids and polypeptide antigens that may be used to formulate vaccines and immune compositions that can be used to immunize and treat persons who are infected with SARS-CoV. In addition, the invention provides antibodies that bind to the spike protein of SARS-CoV which may be used to diagnose, immunize, and treat persons infected with SARS-CoV.

Definitions:

An "adjuvant" is generally defined as a substance that nonspecifically enhances the immune response to an antigen. A variety of adjuvants may be employed with the immunopeptides and immunofragopeptides of this invention. Most adjuvants contain a substance designed to protect the antigen from rapid catabolism, such as aluminum hydroxide or mineral oil, and a stimulator of immune responses, such as lipid A, *Bortadella pertussis* or *Mycobacterium tuberculosis* derived proteins. Suitable adjuvants are commercially available as,

for example, Freund's Incomplete Adjuvant and Complete Adjuvant (Difco Laboratories, Detroit, Mich.); Merck Adjuvant 65 (Merck and Company, Inc., Rahway, N.J.); aluminum salts such as aluminum hydroxide gel (alum) or aluminum phosphate; salts of calcium, iron or zinc; an insoluble suspension of acylated tyrosine; acylated sugars; cationically or anionically derivatized polysaccharides; polyphosphazenes; biodegradable microspheres; monophosphoryl lipid A and quil A. Cytokines, such as GM-CSF or interleukin-2, -7, or -12, may also be used as adjuvants.

An "animal" refers to an organism that can mount an immune response upon antigenic challenge. For example, reptiles, avians, and mammals are able to produce antibodies in response to an antigenic challenge. Antibodies raised in non-human organisms are thought to be useful in diagnostic assays to reduce or eliminate cross-reactivity.

An "aptamer" is a peptide, polypeptide or nucleic acid (RNA or DNA) that binds to a polypeptide or peptide fragment of the invention.

A "carrier protein" refers to a polypeptide that can be coupled with a polypeptide or a peptide fragment of the invention to form a coupled protein. A carrier protein may be coupled to a polypeptide or peptide fragment in order to increase the solubility or the immunogenicity of the polypeptide or peptide fragment. A carrier protein may also be coupled to a polypeptide or peptide fragment to provide a tag which provides for separation or detection of the coupled protein. For example, biotin may be used as a carrier protein that is coupled to a polypeptide or peptide fragment to create a coupled protein which can then be isolated through interaction with avidin, or detected through use of a fluorescently tagged avidin. In another example, a carrier protein that is bound by an antibody can be coupled to a polypeptide or peptide fragment to create a coupled protein that is bound by the antibody which binds to the carrier protein of the coupled protein.

The invention encompasses isolated or substantially purified nucleic acids, peptides, polypeptides or proteins. In the context of the present invention, an "isolated" nucleic acid, DNA or RNA molecule or an "isolated" polypeptide is a nucleic acid, DNA molecule, RNA molecule, or polypeptide that exists apart from its native environment and is therefore not a product of nature. An isolated nucleic acid, DNA molecule, RNA molecule or polypeptide may exist in a

purified form or may exist in a non-native environment such as, for example, a transgenic host cell. A "purified" nucleic acid molecule, peptide, polypeptide or protein, or a fragment thereof, is substantially free of other cellular material, or culture medium when produced by recombinant techniques, or substantially free of chemical precursors or other chemicals when chemically synthesized. In one embodiment, an "isolated" nucleic acid is free of sequences that naturally flank the nucleic acid (*i.e.*, sequences located at the 5' and 3' ends of the nucleic acid) in the genomic DNA of the organism from which the nucleic acid is derived. For example, in various embodiments, the isolated nucleic acid molecule can contain less than about 5 kb, 4 kb, 3 kb, 2 kb, 1 kb, 0.5 kb, or 0.1 kb of nucleotide sequences that naturally flank the nucleic acid molecule in genomic DNA of the cell from which the nucleic acid is derived. A protein, peptide or polypeptide that is substantially free of cellular material includes preparations of protein, peptide or polypeptide having less than about 30%, 20%, 10%, or 5% (by dry weight) of contaminating protein. When the protein of the invention, or biologically active portion thereof, is recombinantly produced, preferably culture medium represents less than about 30%, 20%, 10%, or 5% (by dry weight) of chemical precursors or non-protein-of-interest chemicals.

The terms polypeptide, peptide and protein are used interchangeably herein.

A peptide or polypeptide "fragment" as used herein refers to a less than full length peptide, polypeptide or protein. For example, a peptide or polypeptide fragment can have is at least about 3, at least about 4, at least about 5, at least about 10, at least about 20, at least about 30, at least about 40 amino acids in length, or single unit lengths thereof. For example, fragment may be 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, or more amino acids in length. There is no upper limit to the size of a peptide fragment. However, in some embodiments, peptide fragments can be less than about 500 amino acids, less than about 400 amino acids, less than about 300 amino acids or less than about 250 amino acids in length. Preferably the peptide fragment can elicit an immune response when used to inoculate an animal. A peptide fragment may be used to elicit an immune response by inoculating an animal with a peptide fragment in combination with an adjuvant, a peptide fragment that is coupled to an adjuvant, or a peptide fragment that is coupled to arsanilic acid, sulfanilic acid, an acetyl

group, or a picryl group. A peptide fragment can include a non-amide bond, and can be a peptidomimetic.

The term "soluble" as used herein refers to the ability of a polypeptide to be solvated in an aqueous solution. For example, a soluble peptide can be mixed with an aqueous medium such that at least a detectable portion of the peptide is present in the aqueous medium. The peptide may be detected through use of common techniques, such as absorbance of light, fluorescence, the ability to bind dyes, the ability to reduce silver ions, and the like.

The term "specifically binds" refers to an antibody that binds to a single epitope, but which does not bind to more than one epitope. Accordingly, an antibody that specifically binds to a polypeptide will bind to an epitope that present on the polypeptide, but which is not present on other polypeptides.

I. Polypeptides, peptide fragments, coupled proteins, immunopeptides, and peptidomimetics of the invention

The invention provides a polypeptide which has an amino acid sequence that corresponds to the amino acid sequence of the spike protein from the virus (SARS-CoV) that is etiologically linked to severe acute respiratory syndrome (SARS). A representative amino acid sequence is provided by SEQ ID NO: 1, whose sequence is provided below for easy reference.

```
  1 MFIFLLFLTL TSGSDLDRCT TFDDVQAPNY TQHTSSMRGV
 41 YYPDEIFRSD TLYLTQDLFL PFYSNVTGFH TINHTFGNPV
 81 IPFKDGIYFA ATEKSNVVRG WVFGSTMNNK SQSVIIINNS
121 TNVVIRACNF ELCDNPFFAV SKPMGTQTHT MIFDNAFNCT
161 FEYISDAFSL DVSEKSGNFK HLREFVFKNK DGFLYVYKGY
201 QPIDVVRDLP SGFNTLKPIF KLPLGINITN FRAILTAFSP
241 AQDIWGTSAA AYFVGYLKPT TFMLKYDENG TITDAVDCSQ
281 NPLAELKCSV KSFEIDKGIY QTSNFRVVPS GDVVRFPNIT
321 NLCPFGEVFN ATKFPSVYAW ERKKISNCVA DYSVLYNSTF
361 FSTFKCYGVS ATKLNDLCFS NVYADSFVVK GDDVRQIAPG
401 QTGVIADYNY KLPDDFMGCV LAWNTRNIDA TSTGNYNYKY
441 RYLRHGKLRP FERDISNVPF SPDGKPCTPP ALNCYWPLND
481 YGFYTTTGIG YQPYRVVVLS FELLNAPATV CGPKLSTDLI
521 KNQCVNFNFN GLTGTGVLTP SSKRFQPFQQ FGRDVSDFTD
```

```
 561 SVRDPKTSEI LDISPCAFGG VSVITPGTNA SSEVAVLYQD
 601 VNCTDVSTAI HADQLTPAWR IYSTGNNVFQ TQAGCLIGAE
 641 HVDTSYECDI PIGAGICASY HTVSLLRSTS QKSIVAYTMS
 681 LGADSSIAYS NNTIAIPTNF SISITTEVMP VSMAKTSVDC
 721 NMYICGDSTE CANLLLQYGS FCTQLNRALS GIAAEQDRNT
 761 REVFAQVKQM YKTPTLKYFG GFNFSQILPD PLKPTKRSFI
 801 EDLLFNKVTL ADAGFMKQYG ECLGDINARD LICAQKFNGL
 841 TVLPPLLTDD MIAAYTAALV SGTATAGWTF GAGAALQIPF
 881 AMQMAYRFNG IGVTQNVLYE NQKQIANQFN KAISQIQESL
 921 TTTSTALGKL QDVVNQNAQA LNTLVKQLSS NFGAISSVLN
 961 DILSRLDKVE AEVQIDRLIT GRLQSLQTYV TQQLIRAAEI
1001 RASANLAATK MSECVLGQSK RVDFCGKGYH LMSFPQAAPH
1041 GVVFLHVTYV PSQERNFTTA PAICHEGKAY FPREGVFVFN
1081 GTSWFITQRN FFSPQIITTD NTFVSGNCDV VIGIINNTVY
1121 DPLQPELDSF KEELDKYFKN HTSPDVDLGD ISGINASVVN
1161 IQKEIDRLNE VAKNLNESLI DLQELGKYEQ YIKWPWYVWL
1201 GFIAGLIAIV MVTILLCCMT SCCSCLKGAC SCGSCCKFDE
1241 DDSEPVLKGV KLHYT
```

The invention also provides peptide fragments which have amino acid sequences that correspond to a fragment of the spike protein from the virus (SARS-CoV) that is etiologically linked to severe acute respiratory syndrome (SARS). Such amino acid sequences include those represented by SEQ ID NOs: 13, 14, 15, 20-59, and 61-63. The peptide fragments of SEQ ID NO: 1 can also be three or more amino acids in length, and produce an immune response when used to immunize an animal. These peptide fragments are exemplified by those that are three amino acids in length, or single amino acid units of greater length, such as 4, 5, 6, 7, 8, 9, 10 amino acids in length, and an amino acid sequence that lacks one amino acid from the amino acid sequence corresponding to SEQ ID NO: 1.

The invention also provides coupled proteins having a carrier protein coupled to a polypeptide or peptide fragment of the invention. The carrier protein may be used to increase the solubility of the coupled protein. The carrier protein may also be used to increase the immunogenicity of the coupled protein

to increase production of antibodies that bind to the polypeptide or peptide fragment of the invention. The carrier protein may also be used to provide for the separation or detection of a coupled protein. Accordingly, a coupled protein can be detected or isolated by interaction with other components that bind to the carrier protein portion of the coupled protein. For example, a coupled protein having avidin as a carrier protein can be detected or separated with biotin through use of known methods. Numerous carrier proteins may be used to create coupled proteins of the invention. Examples of such carrier proteins include, keyhole limpet hemacyanin, bovine serum albumin, ovalbumin, mouse serum albumin, rabbit serum albumin, and the like. A carrier protein may be coupled to a polypeptide or peptide fragment of the invention by creation of a fusion protein through use of recombinant methods. A carrier protein may also be coupled to a polypeptide or peptide fragment of the invention through use of chemical linking methods, or through use of a chemical linker. Such coupling methods are known in the art and have been described. Harlow et al., Antibodies: A Laboratory Manual, page 319 (Cold Spring Harbor Pub. 1988); Taylor, Protein Immobilization, Marcel Dekker, Inc., New York, (1991).

The invention provides immunopeptides having a polypeptide or a peptide fragment of the invention coupled to arsanilic acid, sulfanilic acid, an acetyl group, or a picryl group. Methods to couple such groups to peptides are known and have been reported. Weigle, J. Exp. Med., 116:913-928 (1962); Weigle, J. Exp. Med., 122:1049-1062 (1965); Weigle, J. Exp. Med., 121:289-308 (1965).

The polypeptides and peptide fragments of the invention may be in glycosylated form, or in unglycosylated form. A polypeptide or peptide fragment of the invention may be soluble or insoluble in aqueous solution. The polypeptides and peptide fragments of the invention may be conservative variants. A conservative variant is a polypeptide or peptide fragment derived from a full-length polypeptide, such as that exemplified by SEQ ID NO: 1, by deletion (so-called truncation), addition, or subtraction of one or more amino acids to the N-terminal and/or C-terminal end of the full-length polypeptide; deletion, addition or subtraction of one or more amino acids at one or more sites in the full-length polypeptide. Such variants may result from, for example, genetic polymorphism or from human manipulation. Methods for such

manipulations are generally known in the art. For example, amino acid sequence variants of SEQ ID NO: 1 can be prepared by mutagenesis of DNA encoding the polypeptide. Methods for mutagenesis and nucleotide sequence alterations are well known in the art. See, for example, Kunkel, Proc. Natl. Acad. Sci. USA, 82, 488 (1985); Kunkel et al., Methods in Enzymol., 154:367 (1987); U. S. Patent No. 4,873,192; Walker and Gaastra, eds., Techniques in Molecular Biology, MacMillan Publishing Company, New York (1983) and the references cited therein. Guidance as to appropriate amino acid substitutions may be found in the model of Dayhoff et al., Atlas of Protein Sequence and Structure, Natl. Biomed. Res. Found., Washington, C.D. (1978), herein incorporated by reference. Conservative substitutions, such as exchanging one amino acid with another having similar properties, are preferred. For example, substitution of a hydrophobic amino acid for another, or substitution of a hydrophilic amino acid for another. Routine screening assays can be used to determine if a substituted polypeptide or peptide fragment derived from SEQ ID NO: 1 produces an immune response when administered to a mammal. Examples of such screening assays are well known in the art and include enzyme linked immunosorbant assays, radioimmuno assays, chromium release assays, and the like. Such assays have been described. Harlow et al., Antibodies: A Laboratory Manual, page 319 (Cold Spring Harbor Pub. 1988).

The invention provides peptidomimetics of the polypeptides and peptide fragments of the invention. A peptidomimetic describes a peptide analog, such as those commonly used in the pharmaceutical industry as non-peptide drugs, with properties analogous to those of the template peptide. (Fauchere, J., Adv. Drug Res., 15: 29 (1986) and Evans et al., J. Med. Chem., 30:1229 (1987)). Peptidomimetics are structurally similar to polypeptides or peptide fragments having peptide bonds, but have one or more peptide linkages optionally replaced by a linkage such as, $-CH_2NH-$, $-CH_2S-$, $-CH_2-CH_2-$, $-CH=CH-$ (cis and trans), $-COCH_2-$, $-CH(OH)CH_2-$, and $-CH_2SO-$, by methods known in the art. Advantages of peptide mimetics over natural polypeptide embodiments may include more economical production, greater chemical stability, altered specificity and enhanced pharmacological properties such as half-life, absorption, potency and efficacy.

The polypeptides, peptide fragments, coupled proteins, and peptidomimetics of the invention can be modified for in vivo use by the addition, at the amino-terminus and/or the carboxyl-terminus, of a blocking agent to decrease degradation in vivo. This can be useful in those situations in which the polypeptide termini tend to be degraded by proteases prior to cellular uptake. Such blocking agents can include, without limitation, additional related or unrelated peptide sequences that can be attached to the amino and/or carboxyl terminal residues of the polypeptide, peptide fragment, coupled protein, and peptidomimetic to be administered. This can be done either chemically during the synthesis of the polypeptide, peptide fragment, or coupled protein, or by recombinant DNA technology by methods familiar to artisans of average skill. Alternatively, blocking agents such as pyroglutamic acid, or other molecules known in the art, can be attached to the amino and/or carboxyl terminal residues, or the amino group at the amino terminus or carboxyl group at the carboxyl terminus can be replaced with a different moiety. Accordingly, the invention provides polypeptides and peptide fragments that are amino-terminally and carboxyl-terminally blocked.

The ability of a polypeptide or peptide fragment of the invention to produce an immune response may be tested through numerous art recognized methods. For example, for their ability to induce antibody production, or to stimulate a cytotoxic T-lymphocyte response.

The polypeptides and peptide fragments of the invention may be used within screening assays to identify or isolate antibodies that bind to the polypeptides or peptide fragments of the invention, or the spike protein from SARS-CoV. For example, the polypeptides or peptide fragments may be used in phage display assays to isolate antibodies that bind to the polypeptides or peptide fragments. In another example, the polypeptides or peptide fragments of the invention may be bound to a solid support to which antibodies are contacted such that antibodies which bind to the polypeptides or peptide fragments become immobilized on the solid support. These antibodies can be later eluted from the solid support. The polypeptides and peptide fragments of the invention may be used to isolate antibodies according to many other methods known in the art.

Expression systems that may be used for small or large scale production of the, coupled proteins, polypeptides or peptide fragments of the invention

include, but are not limited to, cells or microorganisms that are transformed with a recombinant nucleic acid construct that contains a nucleic acid segment of the invention. Examples of recombinant nucleic acid constructs may include bacteriophage DNA, plasmid DNA, cosmid DNA, or viral expression vectors. Examples of cells and microorganisms that may be transformed include bacteria (for example, E. coli or B. subtilis); yeast (for example, Saccharomyces and Pichia); insect cell systems (for example, baculovirus); plant cell systems; or mammalian cell systems (for example, COS, CHO, BHK, 293, VERO, HeLa, MDCK, W138, and NIH 3T3 cells). Also useful as host cells are primary or secondary cells obtained directly from a mammal that are transfected with a plasmid vector or infected with a viral vector. Examples of suitable expression vectors include, without limitation, plasmids and viral vectors such as herpes viruses, retroviruses, vaccinia viruses, attenuated vaccinia viruses, canary pox viruses, adenoviruses, adeno-associated viruses, lentiviruses and herpes viruses, among others. Synthetic methods may also be used to produce polypeptides and peptide fragments of the invention. Such methods are known and have been reported. Merrifield, Science, 85:2149 (1963).

II. Nucleic acid segments, expression cassettes, and nucleic acid constructs of the invention

The present invention provides isolated nucleic acid segments that encode the polypeptides, peptide fragments, and coupled proteins of the invention. The nucleic acid segments of the invention also include segments that encode for the same amino acids due to the degeneracy of the genetic code. For example, the amino acid threonine is encoded by ACU, ACC, ACA and ACG and is therefore degenerate. It is intended that the invention includes all variations of the polynucleotide segments that encode for the same amino acids. Such mutations are known in the art (Watson et al, Molecular Biology of the Gene, Benjamin Cummings 1987). Mutations also include alteration of a nucleic acid segment to encode for conservative amino acid changes, for example, the substitution of leucine for isoleucine and so forth. Such mutations are also known in the art. Thus, the genes and nucleotide sequences of the invention include both the naturally occurring sequences as well as mutant forms.

The nucleic acid segments of the invention may be contained within a vector. A vector may include, but is not limited to, any plasmid, phagemid, F-factor, virus, cosmid, or phage in double or single stranded linear or circular form which may or may not be self transmissible or mobilizable. The vector can also transform a prokaryotic or eukaryotic host either by integration into the cellular genome or exist extra-chromosomally (e.g. autonomous replicating plasmid with an origin of replication).

Preferably the nucleic acid segment in the vector is under the control of, and operably linked to, an appropriate promoter or other regulatory elements for transcription in vitro or in a host cell, such as a eukaryotic cell, or a microbe, e.g. bacteria. The vector may be a shuttle vector that functions in multiple hosts. The vector may also be a cloning vector that typically contains one or a small number of restriction endonuclease recognition sites at which foreign DNA sequences can be inserted in a determinable fashion. Such insertion can occur without loss of essential biological function of the cloning vector. A cloning vector may also contain a marker gene that is suitable for use in the identification and selection of cells transformed with the cloning vector. Examples of marker genes are tetracycline resistance or ampicillin resistance. Many cloning vectors are commercially available (Stratagene, New England Biolabs, Clonetech).

The nucleic acid segments of the invention may also be inserted into an expression vector. Typically an expression vector contains prokaryotic DNA elements coding for a bacterial replication origin and an antibiotic resistance gene to provide for the amplification and selection of the expression vector in a bacterial host; regulatory elements that control initiation of transcription such as a promoter; and DNA elements that control the processing of transcripts such as introns, or a transcription termination / polyadenylation sequence.

Methods to introduce nucleic acid segment into a vector are available in the art (Sambrook et al., Molecular Cloning: A Laboratory Manual, 3rd edition, Cold Spring Harbor Press, Cold Spring Harbor, N.Y. (2001)). Briefly, a vector into which a nucleic acid segment is to be inserted is treated with one or more restriction enzymes (restriction endonuclease) to produce a linearized vector having a blunt end, a "sticky" end with a 5' or a 3' overhang, or any combination of the above. The vector may also be treated with a restriction enzyme and subsequently treated with another modifying enzyme, such as a polymerase, an

exonuclease, a phosphatase or a kinase, to create a linearized vector that has characteristics useful for ligation of a nucleic acid segment into the vector. The nucleic acid segment that is to be inserted into the vector is treated with one or more restriction enzymes to create a linearized segment having a blunt end, a "sticky" end with a 5' or a 3' overhang, or any combination of the above. The nucleic acid segment may also be treated with a restriction enzyme and subsequently treated with another DNA modifying enzyme. Such DNA modifying enzymes include, but are not limited to, polymerase, exonuclease, phosphatase or a kinase, to create a nucleic acid segment that has characteristics useful for ligation of a nucleic acid segment into the vector.

The treated vector and nucleic acid segment are then ligated together to form a construct containing a nucleic acid segment according to methods available in the art (Sambrook et al., Molecular Cloning: A Laboratory Manual, 3rd edition, Cold Spring Harbor Press, Cold Spring Harbor, N.Y. (2001)). Briefly, the treated nucleic acid fragment and the treated vector are combined in the presence of a suitable buffer and ligase. The mixture is then incubated under appropriate conditions to allow the ligase to ligate the nucleic acid fragment into the vector.

The invention also provides an expression cassette which contains a nucleic acid sequence capable of directing expression of a particular nucleic acid segment of the invention, such as SEQ ID NO: 2, either in vitro or in a host cell. Also, a nucleic acid segment of the invention may be inserted into the expression cassette such that an anti-sense message is produced. The expression cassette is an isolatable unit such that the expression cassette may be in linear form and functional for in vitro transcription and translation assays. The materials and procedures to conduct these assays are commercially available from Promega Corp. (Madison, Wisconsin). For example, an in vitro transcript may be produced by placing a nucleic acid sequence under the control of a T7 promoter and then using T7 RNA polymerase to produce an in vitro transcript. This transcript may then be translated in vitro through use of a rabbit reticulocyte lysate. Alternatively, the expression cassette can be incorporated into a vector allowing for replication and amplification of the expression cassette within a host cell or also in vitro transcription and translation of a nucleic acid segment.

Such an expression cassette may contain one or a plurality of restriction sites allowing for placement of the nucleic acid segment under the regulation of a regulatory sequence. The expression cassette can also contain a termination signal operably linked to the nucleic acid segment as well as regulatory sequences required for proper translation of the nucleic acid segment. The expression cassette containing the nucleic acid segment may be chimeric, meaning that at least one of its components is heterologous with respect to at least one of its other components. The expression cassette may also be one which is naturally occurring but has been obtained in a recombinant form useful for heterologous expression. Expression of the nucleic acid segment in the expression cassette may be under the control of a constitutive promoter or an inducible promoter which initiates transcription only when the host cell is exposed to some particular external stimulus.

The expression cassette may include in the 5'-3' direction of transcription, a transcriptional and translational initiation region, a nucleic acid segment and a transcriptional and translational termination region functional in vivo and /or in vitro. The termination region may be native with the transcriptional initiation region, may be native with the nucleic acid segment, or may be derived from another source.

The regulatory sequence can be a polynucleotide sequence located upstream (5' non-coding sequences), within, or downstream (3' non-coding sequences) of a coding sequence, and which influences the transcription, RNA processing or stability, or translation of the associated coding sequence. Regulatory sequences can include, but are not limited to, enhancers, promoters, repressor binding sites, translation leader sequences, introns, and polyadenylation signal sequences. They may include natural and synthetic sequences as well as sequences which may be a combination of synthetic and natural sequences. While regulatory sequences are not limited to promoters, some useful regulatory sequences include constitutive promoters, inducible promoters, regulated promoters, tissue-specific promoters, viral promoters and synthetic promoters.

A promoter is a nucleotide sequence which controls the expression of the coding sequence by providing the recognition for RNA polymerase and other factors required for proper transcription. A promoter includes a minimal

promoter, consisting only of all basal elements needed for transcription initiation, such as a TATA-box and/or initiator that is a short DNA sequence comprised of a TATA- box and other sequences that serve to specify the site of transcription initiation, to which regulatory elements are added for control of expression. A promoter may be derived entirely from a native gene, or be composed of different elements derived from different promoters found in nature, or even be comprised of synthetic DNA segments. A promoter may contain DNA sequences that are involved in the binding of protein factors which control the effectiveness of transcription initiation in response to physiological or developmental conditions.

The invention also provides a construct containing a vector and an expression cassette. The vector may be selected from, but not limited to, any vector previously described. Into this vector may be inserted an expression cassette through methods known in the art and previously described (Sambrook et al., Molecular Cloning: A Laboratory Manual, 3rd edition, Cold Spring Harbor Press, Cold Spring Harbor, N.Y. (2001)). In one embodiment, the regulatory sequences of the expression cassette may be derived from a source other than the vector into which the expression cassette is inserted. In another embodiment, a construct containing a vector and an expression cassette is formed upon insertion of a nucleic acid segment of the invention into a vector that itself contains regulatory sequences. Thus, an expression cassette is formed upon insertion of the nucleic acid segment into the vector. Vectors containing regulatory sequences are available commercially and methods for their use are known in the art (Clonetech, Promega, Stratagene).

III. Immune compositions and vaccines of the invention

The invention provides immune compositions and vaccines that can be used to produce an immune response against the virus that is etiologically linked to severe acute respiratory syndrome when administered to an animal. The immune response may be a humoral immune response or a cellular immune response.

An immune composition of the invention can include an adjuvant and a nucleic acid, polypeptide, peptide fragment, a peptidomimetic, a coupled protein, an immunopeptide of the invention, or any combination thereof. An immune

composition can contain an adjuvant that is not chemically linked to a polypeptide, peptide fragment, a peptidomimetic, a coupled protein, or an immunopeptide of the invention. An immune composition can contain an adjuvant that is chemically linked to a polypeptide, peptide fragment, a peptidomimetic, a coupled protein, or an immunopeptide of the invention. An immune composition of the invention can also include a pharmaceutically acceptable diluent or carrier.

An immune composition may be manufactured conventionally. In particular, a nucleic acid, polypeptide, peptide fragment, peptidomimetic, coupled protein, immunopeptide, or any combination thereof that is contained in the composition may be combined with a pharmaceutically acceptable diluent or carrier. Examples of pharmaceutically acceptable diluent or carriers include water or a saline solution, such as phosphate-buffered saline (PBS). In general, the pharmaceutically acceptable diluent or carrier is selected on the basis of the mode and route of administration and of standard pharmaceutical practices. Pharmaceutically acceptable diluents and carriers as well as all that is necessary for their use in pharmaceutical compositions are described in Remington's Pharmaceutical Sciences, a standard reference text in this field.

Immune compositions may contain adjuvants as disclosed herein and as known in the art. Aluminum compounds may be used as adjuvants. Such aluminum compounds include, aluminum hydroxide, aluminum phosphate, aluminum hydroxyphosphate, and the like. The nucleic acid, polypeptide, peptide fragment, peptidomimetic, coupled protein, immunopeptide, or any combination thereof may be absorbed or precipitated on an aluminum compound according to standard methods. Other adjuvants include polyphosphazene (WO 95/2415), DC-chol (3-beta-[N-(N', N'-dimethylaminomethane) carbamoyl) cholesterol] (U.S. Pat. No. 5,283,185 and WO 96/14831), QS-21 (WO 88/9336) and RIBI from ImmunoChem (Hamilton, Montana). Immunostimulatory oligonucleotides containing unmethylated CpG dinucleotides ("CpG") are known in the art as being adjuvants when administered by both systemic and mucosal routes (WO 96/02555, EP 468520, Davis et al., J. Immunol., 160:870 (1998); McCluskie and Davis, J. Immunol., 161:4463 (1998). CpG when formulated into immune compositions or vaccines, is generally administered in free solution together with free antigen (WO 96/02555; McCluskie and Davis, J.

Immunol., 161:4463 (1998)) or covalently conjugated to an antigen (PCT Publication No. WO 98/16247), or formulated with a carrier such as aluminum hydroxide. (Brazolot-Millan et al., Proc.Natl.Acad.Sci., 95:15553 (1998)).

The invention also provides vaccines that include a nucleic acid, polypeptide, a peptide fragment, a peptidomimetic, a coupled protein, an immunopeptide of the invention, a nucleic or any combination thereof. Such vaccines can be formulated as described herein or as known in the vaccine arts. For example, a viral vaccine may be created that expresses a polypeptide, a peptide fragment, or a coupled protein of the invention according to methods known in the art. Examples of viral vectors that may be used include, adenoviruses, herpes viruses, vaccinia viruses, canarypox viruses, and the like. Vaccines can also be formulated as a liposome. Such formulations are known to those skilled in the art. Liposomes: A Practical Approach. RRC New Ed, IRL press (1990).

The invention also provides nucleic acid based vaccines that express a polypeptide, a peptide fragment, or a coupled protein of the invention. For example, a nucleic acid vaccine can express a polypeptide having SEQ ID NO: 1, 13, 14, 15, 20-59, 61-63 or a fragment of SEQ ID NO: 1. Inoculation of an animal with a nucleic acid construct that encodes a polypeptide, a peptide fragment, or a coupled protein of the invention may lead to a humoral and cell-mediated immune response to the encoded antigen. It is thought that some bone marrow-derived professional antigen presenting cells are transfected by the nucleic acid construct and the encoded antigen is transcribed and translated into an immunogenic polypeptide that elicits specific responses. A feature of nucleic acid vaccines is that they provide for eliciting strong cytotoxic T-lymphocyte (CTL) responses. These responses occur because the nucleic acid-encoded polypeptides are synthesized in the cytosol of transfected cells. Furthermore, nucleic acid constructs that are produced in bacteria are rich in unmethylated CpG nucleotides that are recognized as foreign by macrophages. Thus, they elicit an innate immune response that enhances adaptive immunity. Therefore, nucleic acid vaccines are effective even when administered without adjuvants.

Direct injection of an expression cassette into living host cells transforms a number of the cells and causes them to express the introduced nucleic acid and thereby express a gene product. The transfected cells may display fragments of

the expressed antigens on their cell surfaces together with major histocompatibility class I (MHC I) or class II (MHC II) complexes.

Nucleic acid constructs can be introduced into cells more efficiently by inducing muscle degeneration prior to the injection of the nucleic acid construct into an animal, including a human (Vitadello et. al., Hum. Gene. Ther., 5:11 (1994) ; Danko and Wolff, Vaccine, 12:1499 (1994); Davis et. al., Hum. Gene. Ther., 4:733 (1993)). For example, such a treatment is thought to increase the efficiency of transfer by up to 40 fold. Two of the most commonly used myonecrotic agents are the local anesthetic bupivicaine, and cardiotoxin (Danko and Wolff, Vaccine, 12:1499 (1994); Davis et. al., Hum. Gene. Ther., 4:733 (1993)). A number of other techniques have been employed to transfer nucleic acid constructs to muscle. Such other techniques include retroviral vectors, adenoviral vectors, and liposomes. However, direct injection of naked nucleic acid appears to be the most efficient of these delivery mechanisms at transferring and expressing foreign nucleic acids in cells.

Nucleic acid constructs can be administered in a pharmaceutically acceptable carrier. Pharmaceutically acceptable carriers are biologically compatible vehicles which are suitable for administration to a human or other mammalian subject, e.g., physiological saline. A therapeutically effective amount is an amount of the nucleic acid construct that is capable of producing an immune response (e.g., an enhanced T-cell response or antibody production) in a treated animal. As is well known in the medical arts, the dosage for any one patient depends upon many factors, including the patient's size, body surface area, age, the particular compound to be administered, sex, time and route of administration, general health, and other drugs being administered concurrently. Dosages will vary, but a preferred dosage for administration of a nucleic acid construct is from approximately 10^6 to 10^{12} copies of the nucleic acid construct. This does can be repeatedly administered, as needed.

Numerous routes of administration may be used to administer nucleic acid constructs. Examples of such routes include intramuscular injection, intravenous, intraperitoneal, intradermal, intranasal and subcutaneous injection of nucleic acid constructs have all resulted in immunization against influenza virus hemagglutinin (HA) in chickens (reviewed in Pardoll and Beckerkleg, Immunity 3 (1995), 165-169). Nucleic acid based vaccines can also be

administered through use of a polymeric, biodegradable microparticle or microcapsule delivery vehicle, sized to optimize phagocytosis by phagocytic cells such as macrophages. For example, PLGA (poly-lacto-co-glycolide) microparticles approximately 1-10 μm in diameter can be used. The nucleic acid construct is encapsulated in these microparticles, which are taken up by macrophages and gradually biodegraded within the cell, thereby releasing the nucleic acid construct. Once released, the nucleic acid is expressed within the cell. Another way to achieve uptake of a nucleic acid construct is through use of liposomes. Such liposomes can be prepared by standard methods. The nucleic acid constructs can be incorporated alone into these delivery vehicles or co-incorporated with tissue-specific antibodies. Alternatively, a molecular conjugate can be prepared that is composed of a nucleic acid construct attached to poly-L-lysine by electrostatic or covalent forces. Poly-L-lysine binds to a ligand that can bind to a receptor on target cells. Cristiano et al. (1995), J. Mol. Med. 73, 479). Alternatively, lymphoid tissue specific targeting can be achieved by the use of lymphoid tissue-specific transcriptional regulatory elements (TRE) such as a B lymphocyte, T lymphocyte, or dendritic cell specific TRE. Lymphoid tissue specific TRE are known (Thompson et al., Mol. Cell. Biol., 12:1043 (1992); Todd et al., J. Exp. Med., 177:1663 (1993); Penix et al., J. Exp. Med., 178:1483 (1993)).

The invention also provides microbe based vaccines. Generally, these vaccines relate to microbes that have been transformed with a nucleic acid construct that provides for the expression of a polypeptide, a peptide fragment, or a coupled protein of the invention. For example, Listeria monocytogenes may be used as a vector to elicit T-cell immunity. This is because it infects antigen-presenting cells and also because infection originates at the mucosa. Lieberman and Frankel, Vaccine, 20:2007-10 (2002). According, Listeria may be transformed with a nucleic acid construct that provides for the expression of a polypeptide, a peptide fragment, or a coupled protein that elicits an immune response against the spike protein from the coronavirus that causes severe acute respiratory syndrome. Highly attenuated forms of Listeria may be constructed according to methods reported in the art. Lieberman and Frankel, Vaccine, 20:2007 (2002). Salmonella may also be used as a vector to elicit a cytotoxic T

lymphocyte (CTL) response against the coronavirus that causes severe acute respiratory syndrome. Pasetti et al., Infect Immun., 70:4009 (2002).

An immune composition or vaccine may be administered by any conventional route used in the field of vaccines. For example, an immune composition or vaccine can be administered orally or by intravenous infusion, or injected subcutaneously, intramuscularly, intraperitoneally, intrarectally, intravaginally, intranasally, intragastrically, intratracheally, or intrapulmonarily. The choice of the administration route depends on a number of parameters such as the nature of the active principle; the identity of the polypeptide, peptide fragment, peptidomimetic, coupled protein, immunopeptide, DNA vaccine; or the adjuvant that is combined with the aforementioned molecules. Administration of an immune composition may take place in a single dose or in a dose repeated once or several times over a certain period. The appropriate dosage varies according to various parameters. Such parameters include the individual treated (adult or child), the immune composition or antigen itself, the mode and frequency of administration, the presence or absence of adjuvant and, if present, the type of adjuvant and the desired effect (e.g. protection or treatment), as will be determined by persons skilled in the art.

IV. Antibodies and aptamers of the invention

The invention provides antibodies that bind to an amino acid sequence as set forth in SEQ ID NO: 1, 13, 14, 15, 20-59, 60, 61, 62, 63 or a fragment of SEQ ID NO: 1, or conservative variants thereof. Such antibodies are useful for the diagnosis, immunization against, and treatment of severe acute respiratory syndrome (SARS). In some embodiments, the antibody binds to a peptide having SEQ ID NO:58 or 59. Antibodies that bind to the P540 peptide (SEQ ID NO:59) are highly effective, and can detect spike polypeptides even after extensive dilution. For example, a P540 antibody preparation diluted 1:10,000 could still detect spike polypeptides.

Antibodies can be prepared using an intact polypeptide or peptide fragment of interest as the immunizing antigen. The polypeptide or fragment used to immunize an animal can be derived from translated cDNA or chemical synthesis. A polypeptide or peptide fragment can be coupled to a carrier protein, if desired. Such commonly used carrier proteins which are chemically coupled

to the peptide include keyhole limpet hemocyanin (KLH), thyroglobulin, bovine serum albumin (BSA), and tetanus toxoid. A coupled protein can be used to immunize the animal (e.g., a mouse, a rat, or a rabbit).

If desired, polyclonal or monoclonal antibodies can be further purified, for example, by binding to and elution from a matrix to which the polypeptide or peptide fragment to which the antibodies were raised is bound. Those of skill in the art will know of various techniques common in the immunology arts for purification and/or concentration of polyclonal antibodies, as well as monoclonal antibodies (Coligan, et al., Unit 9, Current Protocols in Immunology, Wiley Interscience, 1991, incorporated by reference).

It is also possible to use the anti-idiotype technology to produce monoclonal antibodies which mimic an epitope. For example, an anti-idiotypic monoclonal antibody made to a first monoclonal antibody will have a binding domain in the hypervariable region which is the "image" of the epitope bound by the first monoclonal antibody.

An antibody suitable for binding to a polypeptide or peptide fragment is specific for at least one portion of a region of the polypeptide. For example, one of skill in the art can use a peptide fragment to generate appropriate antibodies of the invention. Antibodies of the invention include polyclonal antibodies, monoclonal antibodies, and fragments of polyclonal and monoclonal antibodies.

The preparation of polyclonal antibodies is well-known to those skilled in the art (Green et al., Production of Polyclonal Antisera, in Immunochemical Protocols (Manson, ed.), pages 1-5 (Humana Press 1992); Coligan et al., Production of Polyclonal Antisera in Rabbits, Rats, Mice and Hamsters, in Current Protocols in Immunology, section 2.4.1 (1992), which are hereby incorporated by reference). For example, a polypeptide or peptide fragment is injected into an animal host, preferably according to a predetermined schedule incorporating one or more booster immunizations, and the animal is bled periodically. Polyclonal antibodies specific for the polypeptide or peptide fragment may then be purified from such antisera by, for example, affinity chromatography using the polypeptide or peptide fragment coupled to a suitable solid support.

The preparation of monoclonal antibodies likewise is conventional (Kohler & Milstein, Nature, 256:495 (1975); Coligan et al., sections 2.5.1-2.6.7;

and Harlow et al., Antibodies: A Laboratory Manual, page 726 (Cold Spring Harbor Pub. 1988)), which are hereby incorporated by reference. Briefly, monoclonal antibodies can be obtained by injecting mice with a composition comprising an antigen, verifying the presence of antibody production by removing a serum sample, removing the spleen to obtain B lymphocytes, fusing the B lymphocytes with myeloma cells to produce hybridomas, cloning the hybridomas, selecting positive clones that produce antibodies to the antigen, and isolating the antibodies from the hybridoma cultures. Monoclonal antibodies can be isolated and purified from hybridoma cultures by a variety of well-established techniques. Such isolation techniques include affinity chromatography with Protein-A Sepharose, size-exclusion chromatography, and ion-exchange chromatography (Coligan et al., sections 2.7.1-2.7.12 and sections 2.9.1-2.9.3; Barnes et al., Purification of Immunoglobulin G (IgG), in Methods in Molecular Biology, Vol. 10, pages 79-104 (Humana Press 1992)). Methods of in vitro and in vivo multiplication of monoclonal antibodies is well-known to those skilled in the art. Multiplication in vitro may be carried out in suitable culture media such as Dulbecco's Modified Eagle Medium or RPMI 1640 medium, optionally replenished by a mammalian serum such as fetal calf serum or trace elements and growth-sustaining supplements such as normal mouse peritoneal exudate cells, spleen cells, bone marrow macrophages. Production in vitro provides relatively pure antibody preparations and allows scale-up to yield large amounts of the desired antibodies. Large scale hybridoma cultivation can be carried out by homogenous suspension culture in an air reactor, in a continuous stirrer reactor, or immobilized or entrapped cell culture. Multiplication in vivo may be carried out by injecting cell clones into mammals histocompatible with the parent cells, e.g., osyngeneic mice, to cause growth of antibody-producing tumors. Optionally, the animals are primed with a hydrocarbon, especially oils such as pristine tetramethylpentadecane prior to injection. After one to three weeks, the desired monoclonal antibody is recovered from the body fluid of the animal.

Antibodies can also be prepared through use of phage display techniques. In one example, an organism is immunized with an antigen, such as a polypeptide or peptide fragment of the invention. Lymphocytes are isolated from the spleen of the immunized organism. Total RNA is isolated from the

splenocytes and mRNA contained within the total RNA is reverse transcribed into complementary deoxyribonucleic acid (cDNA). The cDNA encoding the variable regions of the light and heavy chains of the immunoglobulin is amplified by polymerase chain reaction (PCR). To generate a single chain fragment variable (scFV) antibody, the light and heavy chain amplification products may be linked by splice overlap extension PCR to generate a complete sequence and ligated into a suitable vector. E. coli are then transformed with the vector encoding the scFV, and are infected with helper phage, to produce phage particles that display the antibody on their surface. Alternatively, to generate a complete antigen binding fragment (Fab), the heavy chain amplification product can be fused with a nucleic acid sequence encoding a phage coat protein, and the light chain amplification product can be cloned into a suitable vector. E. coli expressing the heavy chain fused to a phage coat protein are transformed with the vector encoding the light chain amplification product. The disulphide linkage between the light and heavy chains are established in the periplasm of E. coli. The result of this procedure is to produce an antibody library with up to 10^9 clones. The size of the library can be increased to 10^{18} phage by later addition of the immune responses of additional immunized organisms that may be from the same or different hosts. Antibodies that recognize a specific antigen can be selected through panning. Briefly, an entire antibody library can be exposed to an immobilized antigen against which antibodies are desired. Phage that do not express an antibody that binds to the antigen are washed away. Phage that express the desired antibodies are immobilized on the antigen. These phage are then eluted and again amplified in E. coli. This process can be repeated to enrich the population of phage that express antibodies that specifically bind to the antigen. After phage are isolated that express an antibody that binds to an antigen, a vector containing the coding sequences for the antibody can be isolated from the phage particles and the coding sequences can be recloned into a suitable vector to produce an antibody in soluble form. In another example, a human phage library can be used to select for antibodies, such as monoclonal antibodies, that bind to the spike protein from SARS-CoV. Briefly, splenocytes may be isolated from a human that is infected, or not infected, with SARS-CoV and used to create a human phage library according to methods as described above and known in the art. These methods may be used to obtain human

monoclonal antibodies that bind to the spike protein of SARS-CoV. Phage display methods to isolate antigens and antibodies are known in the art and have been described (Gram et al., Proc. Natl. Acad. Sci., 89:3576 (1992); Kay et al., Phage display of peptides and proteins: A laboratory manual. San Diego: Academic Press (1996); Kermani et al., Hybrid, 14:323 (1995); Schmitz et al., Placenta, 21 Suppl. A:S106 (2000); Sanna et al., Proc. Natl. Acad. Sci., 92:6439 (1995)).

An antibody of the invention may be derived from a "humanized" monoclonal antibody. Humanized monoclonal antibodies are produced by transferring mouse complementarity determining regions from heavy and light variable chains of the mouse immunoglobulin into a human variable domain, and then substituting human residues in the framework regions of the murine counterparts. The use of antibody components derived from humanized monoclonal antibodies obviates potential problems associated with the immunogenicity of murine constant regions. General techniques for cloning murine immunoglobulin variable domains are described (Orlandi et al., Proc. Nat'l Acad. Sci. USA, 86:3833 (1989) which is hereby incorporated in its entirety by reference). Techniques for producing humanized monoclonal antibodies are described (Jones et al., Nature, 321:522 (1986); Riechmann et al., Nature, 332:323 (1988); Verhoeyen et al, Science, 239:1534 (1988); Carter et al., Proc. Nat'l Acad. Sci. USA, 89:4285 (1992); Sandhu, Crit. Rev. Biotech., 12:437 (1992); and Singer et al., J. Immunol., 150:2844 (1993), which are hereby incorporated by reference).

In addition, antibodies of the present invention may be derived from a human monoclonal antibody. Such antibodies are obtained from transgenic mice that have been "engineered" to produce specific human antibodies in response to antigenic challenge. In this technique, elements of the human heavy and light chain loci are introduced into strains of mice derived from embryonic stem cell lines that contain targeted disruptions of the endogenous heavy and light chain loci. The transgenic mice can synthesize human antibodies specific for human antigens, and the mice can be used to produce human antibody-secreting hybridomas. Methods for obtaining human antibodies from transgenic mice are described (Green et al., Nature Genet., 7:13 (1994); Lonberg et al., Nature,

368:856 (1994); and Taylor et al., Int. Immunol., 6:579 (1994), which are hereby incorporated by reference).

Antibody fragments of the invention can be prepared by proteolytic hydrolysis of the antibody or by expression in E. coli of DNA encoding the fragment. Antibody fragments can be obtained by pepsin or papain digestion of whole antibodies by conventional methods. For example, antibody fragments can be produced by enzymatic cleavage of antibodies with pepsin to provide a 5S fragment denoted $F(ab')_2$. This fragment can be further cleaved using a thiol reducing agent, and optionally a blocking group for the sulfhydryl groups resulting from cleavage of disulfide linkages, to produce 3.5S Fab' monovalent fragments. Alternatively, an enzymatic cleavage using pepsin produces two monovalent Fab' fragments and an Fc fragment directly. These methods are described (U.S. patents No. 4,036,945; 4,331,647; and 6,342,221, and references contained therein; Porter, Biochem. J., 73:119 (1959); Edelman et al., Methods in Enzymology, Vol. 1, page 422 (Academic Press 1967); and Coligan et al. at sections 2.8.1-2.8.10 and 2.10.1-2.10.4).

Other methods of cleaving antibodies, such as separation of heavy chains to form monovalent light-heavy chain fragments, further cleavage of fragments, or other enzymatic, chemical, or genetic techniques may also be used, so long as the fragments bind to the antigen that is recognized by the intact antibody.

For example, Fv fragments comprise, an association of V_H and V_L chains. This association may be noncovalent (Inbar et al., Proc. Nat'l Acad. Sci. USA, 69:2659 (1972)). Alternatively, the variable chains can be linked by an intermolecular disulfide bond or cross-linked by chemicals such as glutaraldehyde (Sandhu, Crit. Rev. Biotech., 12:437 (1992)). Preferably, the Fv fragments comprise V_H and V_L chains connected by a peptide linker. These single-chain antigen binding proteins (sFv) are prepared by constructing a structural gene comprising DNA sequences encoding the V_H and V_L domains connected by an oligonucleotide. The structural gene is inserted into an expression vector, which is subsequently introduced into a host cell such as E. coli. The recombinant host cells synthesize a single polypeptide chain with a linker peptide bridging the two V domains. Methods for producing sFvs are described (Whitlow et al., Methods: A Companion to Methods in Enzymology, Vol. 2, page 97 (1991); Bird et al., Science, 242:423 (1988), Ladner et al., U.S.

patent No. 4,946,778; Pack et al., Bio/Technology, 11:1271 (1993); and Sandhu, Crit. Rev. Biotech., 12:437 (1992)).

Another form of an antibody fragment is a peptide that forms a single complementarity-determining region (CDR). CDR peptides ("minimal recognition units") can be obtained by constructing genes encoding the CDR of an antibody of interest. Such genes are prepared, for example, by using the polymerase chain reaction to synthesize the variable region from RNA of antibody-producing cells (Larrick et al., Methods: A Companion to Methods in Enzymology, Vol. 2, page 106 (1991)).

An antibody of the invention may be coupled to a toxin. Such antibodies may be used to treat animals, including humans, that are infected with the virus that is etiologically linked to severe acute respiratory syndrome. For example, an antibody that binds to the spike protein of the coronavirus that is etiologically linked to severe acute respiratory syndrome may be coupled to a tetanus toxin and administered to an animal suffering from infection by the aforementioned virus. The toxin-coupled antibody is thought to bind to a portion of a spike protein presented on an infected cell, and then kill the infected cell.

An antibody of the invention may be coupled to a detectable tag. Such antibodies may be used within diagnostic assays to determine if an animal, such as a human, is infected with SARS-CoV. Examples of detectable tags include, fluorescent proteins (i.e., green fluorescent protein, red fluorescent protein, yellow fluorescent protein), fluorescent markers (i.e., fluorescein isothiocyanate, rhodamine, texas red), radiolabels (i.e., ^{3}H, ^{32}P, ^{125}I), enzymes (i.e., β-galactosidase, horseradish peroxidase, β-glucuronidase, alkaline phosphatase), or an affinity tag (i.e., avidin, biotin, streptavidin). Methods to couple antibodies to a detectable tag are known in the art. Harlow et al., Antibodies: A Laboratory Manual, page 319 (Cold Spring Harbor Pub. 1988).

The invention also provides aptamers to the polypeptides and peptide fragments of the invention. Aptamers of the invention can be peptide or nucleic acid aptamers. Peptide aptamers are peptides that bind to a polypeptide or peptide fragment of the invention with affinities that are often comparable to those for monoclonal antibody-antigen complexes. Similarly, nucleic acid aptamers are nucleic acids that bind to a polypeptide or peptide fragment of the

invention with strong affinities, for example, affinities that are often comparable to those for monoclonal antibody-antigen complexes.

In one example, nucleic acid aptamers can be isolated through use of a library of random oligonucleotide sequences. The library is screened to ascertain which oligonucleotide binds to the S polypeptides and peptide fragments of the invention. The bound oligonucleotides are eluted from the immobilized polypeptides or peptide fragments and are then amplified by PCR. This process may be repeated to select for aptamers having high affinity for the polypeptides and peptide fragments of the invention. The sequence of the nucleic acid coding for the aptamers can then be determined and cloned into a suitable vector to facilitate production and maintenance of the desired aptamer..

Peptide aptamers can be isolated by mRNA display of a library that contains a promoter, a start codon, a nucleic acid sequence that encodes random peptides. In some embodiments, the DNA library also includes a nucleic acid segment that codes for a histidine tag. This library is transcribed using a suitable polymerase, such as T7 RNA polymerase, after which a puromycin-containing poly A linker is ligated onto the 3' end of the newly formed mRNAs. When these mRNAs are translated *in vitro,* the nascent peptides form covalent bonds to the puromycin of the linker to form an mRNA-peptide fusion molecule. The mRNA-peptide fusion molecules are then purified through use of Ni-NTA agarose and oligo-dT-cellulose. The mRNA portion of the fusion molecule is then reverse transcribed. The double-stranded DNA/RNA-peptide fusion molecules are then incubated with a polypeptide or peptide fragment of the invention and unbound fusion molecules are washed away. The bound fusion molecules are eluted from the immobilized polypeptides or peptide fragments and are then amplified by PCR. This process may be repeated to select for aptamers having high affinity for the polypeptides and peptide fragments of the invention. The sequence of the nucleic acid coding for the aptamers can then be determined and cloned into a suitable vector. Methods for the preparation of peptide aptamers have been described (Wilson et al., Proc. Natl. Acad. Sci., 98:3750 (2001)). Accordingly, the invention provides aptamers that recognize the polypeptides and peptide fragments of the invention.

V. Pharmaceutical compositions of the invention

The invention provides pharmaceutical compositions containing an antibody that binds to an amino acid sequence as set forth in SEQ ID NO: 1, 13, 14, 15, 20-59, 60, 61, 62, 63 or a fragment of SEQ ID NO: 1, or a conservative variant thereof, and a pharmaceutically acceptable carrier. In some embodiments, the antibody binds to a peptide having SEQ ID NO:58 or 59. Antibodies that bind to the P540 peptide (SEQ ID NO:59) are highly effective, and can detect spike polypeptides even after extensive dilution. For example, a P540 antibody preparation at dilution 1:10,000 could still detect spike polypeptides.

The pharmaceutical compositions of the invention may be prepared in many forms that include tablets, hard or soft gelatin capsules, aqueous solutions, suspensions, and liposomes and other slow-release formulations, such as shaped polymeric gels. An oral dosage form may be formulated such that the antibody is released into the intestine after passing through the stomach. Such formulations are described in U.S. Patent No. 6,306,434 and in the references contained therein.

Oral liquid pharmaceutical compositions may be in the form of, for example, aqueous or oily suspensions, solutions, emulsions, syrups or elixirs, or may be presented as a dry product for constitution with water or other suitable vehicle before use. Such liquid pharmaceutical compositions may contain conventional additives such as suspending agents, emulsifying agents, non-aqueous vehicles (which may include edible oils), or preservatives.

An antibody can be formulated for parenteral administration (e.g., by injection, for example, bolus injection or continuous infusion) and may be presented in unit dosage form in ampules, prefilled syringes, small volume infusion containers or multi-dose containers with an added preservative. The pharmaceutical compositions may take such forms as suspensions, solutions, or emulsions in oily or aqueous vehicles, and may contain formulatory agents such as suspending, stabilizing and/or dispersing agents. Pharmaceutical compositions suitable for rectal administration can be prepared as unit dose suppositories. Suitable carriers include saline solution and other materials commonly used in the art.

For administration by inhalation, an antibody can be conveniently delivered from an insufflator, nebulizer or a pressurized pack or other convenient

means of delivering an aerosol spray. Pressurized packs may comprise a suitable propellant such as dichlorodifluoromethane, trichlorofluoromethane, dichlorotetrafluoroethane, carbon dioxide or other suitable gas. In the case of a pressurized aerosol, the dosage unit may be determined by providing a valve to deliver a metered amount.

Alternatively, for administration by inhalation or insufflation, an antibody may take the form of a dry powder composition, for example, a powder mix of a modulator and a suitable powder base such as lactose or starch. The powder composition may be presented in unit dosage form in, for example, capsules or cartridges or, e.g., gelatin or blister packs from which the powder may be administered with the aid of an inhalator or insufflator. For intra-nasal administration, an antibody may be administered via a liquid spray, such as via a plastic bottle atomizer.

Pharmaceutical compositions of the invention may also contain other ingredients such as flavorings, colorings, anti-microbial agents, or preservatives. It will be appreciated that the amount of an antibody required for use in treatment will vary not only with the particular carrier selected but also with the route of administration, the nature of the condition being treated and the age and condition of the patient. Ultimately the attendant health care provider may determine proper dosage. In addition, a pharmaceutical composition may be formulated as a single unit dosage form.

VI. <u>Method to immunize, treat, and diagnose an animal against severe acute respiratory syndrome</u>

The invention provides a method to immunize an animal against severe acute respiratory syndrome. The method relates to administering a therapeutically effective amount of an antibody that binds to an amino acid sequence as set forth in SEQ ID NO: 1, 13, 14, 15, 20-59, 60, 61, 62, 63 or a fragment of SEQ ID NO: 1, or a conservative variant thereof to an animal; administering an effective amount of an immune composition to an animal; administering an effective amount of a viral vaccine to an animal; or administering an effective amount of a nucleic acid vaccine to an animal. The animal may be a mammal, such as a human. Methods to administer vaccines and immune compositions have been described herein and are known in the art.

An animal may also be treated for infection by SARS-CoV through passive immunization according to the invention. For example, antibodies that bind to an amino acid sequence as set forth in SEQ ID NO: 1, 13, 14, 15, 20-55, 60, 61, 62, 63 or a fragment of SEQ ID NO: 1, or a conservative variant thereof may be administered to an animal, such as a human, that is infected with SARS-CoV. Such administration may be suitable in situations where a patient is immune compromised and is unable to mount an effective immune response against SARS-CoV, or to a vaccine or immune composition.

The invention provides a method to diagnose severe acute respiratory syndrome in an animal that involves contacting a biological sample obtained from the animal, such as tissue samples, blood, mucus, or saliva, with an antibody that binds to an amino acid sequence as set forth in SEQ ID NO: 1, 13, 14, 15, 20-59, 60, 61, 62, 33 or a fragment of SEQ ID NO: 1, and determining if the antibody binds to the biological sample. Diagnostic assays that utilize antibodies to detect the presence of an antigen in a biological sample are well known in the art. Briefly, an antibody of the invention may be immobilized on a surface. A biological sample can then be contacted with the immobilized antibody such that an antigen contained in the sample is bound by the antibody to form an antibody-antigen complex. The sample may then be optionally washed to remove unbound materials. A second antibody of the invention that is coupled to a detectable tag, such as an enzyme or radiolabel, can then be contacted with the antibody-antigen complex such that the enzyme or radiolabel is immobilized on the surface. The detectable tag can then be detected to determine if an antigen was present in the biological sample. In another example, a biological sample can be immobilized on a surface. An antibody of the invention that is coupled to a detectable tag is then contacted with the immobilized biological sample and any unbound material is washed away. The presence of the detectable tag is then detected to determine whether the biological sample contained an antigen. Examples of such assays are well known in the art and include, enzyme-linked immunosorbant assays, radioimmuno assays, and the like.

Nucleic acid based methods may also be used to diagnose severe acute respiratory syndrome. In one example, polymerase chain reaction (PCR) may be used to diagnose SARS-CoV infection. Briefly, a biological sample, such as a

tissue sample, blood, mucus, or saliva, is obtained from an animal. The nucleic acids within the sample are then extracted using common methods, such as organic extraction. The extracted nucleic acids are then mixed with forward and reverse primers that anneal to nucleic acids that encode SARS proteins, polymerase, nucleotides, and typically a buffer that includes components that allow the polymerase to extend the forward and reverse primers using the SARS nucleic acid as a template. The presence of amplified DNA between the forward and reverse primers is then detected to determine if the sample contained SARS originated nucleic acid. Nucleic acid hybridization techniques, such as Northern and Southern blotting, may also be used to detect the presence of SARS nucleic acids in a biological sample.

VII. Kits

The invention provides a kit which contains packaging material and an antibody that binds to an amino acid sequence as set forth in SEQ ID NO: 1, 13, 14, 15, 45, 46, or 47, 58, 59, 61, 62, 63 or a fragment of SEQ ID NO: 1, or a conservative variant thereof. The kit may also contain a syringe to allow for injection of the antibody contained within the kit into an animal, such as a human. In another embodiment, the invention provides a kit that may contain packaging material, and an antibody that binds to an amino acid sequence as set forth in SEQ ID NO: 1, 13, 14, 15, 20-59, 60, 61, 62, 63 or a fragment of SEQ ID NO: 1, or a conservative variant thereof that is formulated for administration to an animal, such as a human. In some embodiments, the antibody binds to an amino acid sequence set forth in SEQ ID NO:59. In other embodiments, the antibody binds to an amino acid sequence as set forth in SEQ ID NO:58. Such a kit may optionally contain a syringe to allow for injection of the antibody contained within the kit into an animal, such as a human.

The invention also provides a kit which contains packaging material and DNA vaccine having a DNA molecule or expression vector encoding a polypeptide with an amino acid sequence as set forth in SEQ ID NO: 1, 13, 14, 15, 45, 46, or 47, 58, 59, 61, 62, 63 or a fragment of SEQ ID NO: 1, or a conservative variant thereof. The kit may also contain a device for administering the DNA vaccine (e.g. a syringe or gene gun) to allow for administration of the vaccine contained within the kit into an animal, such as a human.

The invention also provides a kit which contains packaging material and vaccine composition that includes a polypeptide with an amino acid sequence as set forth in SEQ ID NO: 1, 13, 14, 15, 45, 46, or 47, 58, 59, 61, 62, 63 or a fragment of SEQ ID NO: 1, or a conservative variant thereof. The kit may also contain a device for administering the vaccine (e.g. a syringe) to allow for administration of the vaccine contained within the kit into an animal, such as a human.

The invention also provides a kit for detecting SARS-CoV infection, which contains packaging material and a polypeptide with an amino acid sequence as set forth in SEQ ID NO: 1, 13, 14, 15, 45, 46, or 47, 58, 59, 61, 62, 63 or a fragment of SEQ ID NO: 1, or a conservative variant thereof. The polypeptide(s) can be immobilized onto a solid support. Such a kit may be used for detection of antibodies directed against the SARS-CoV in the serum of infected animals or humans. The kit can also contain a means for detecting binding of such antibodies to the S polypeptide(s).

VIII. <u>Amino Acid sequence of a full-length spike (S) protein (amino acids 1-1255) from the Tor2 isolate of the SARS-CoV virus</u>

MFIFLLFLTLTSGSDLDRCTTFDDVQAPNYTQHTSSMRGVYYPDEIFRSD
TLYLTQDLFLPFYSNVTGFHTINHTFGNPVIPFKDGIYFAATEKSNVVRG
WVFGSTMNNKSQSVIIINNSTNVVIRACNFELCDNPFFAVSKPMGTQTHT
MIFDNAFNCTFEYISDAFSLDVSEKSGNFKHLREFVFKNKDGFLYVYKG
YQPIDVVRDLPSGFNTLKPIFKLPLGINITNFRAILTAFSPAQDIWGTSAAA
YFVGYLKPTTFMLKYDENGTITDAVDCSQNPLAELKCSVKSFEIDKGIYQ
TSNFRVVPSGDVVRFPNITNLCPFGEVFNATKFPSVYAWERKKISNCVAD
YSVLYNSTFFSTFKCYGVSATKLNDLCFSNVYADSFVVKGDDVRQIAPG
QTGVIADYNYKLPDDFMGCVLAWNTRNIDATSTGNYNYKYRYLRHGK
LRPFERDISNVPFSPDGKPCTPPALNCYWPLNDYGFYTTTGIGYQPYRVV
VLSFELLNAPATVCGPKLSTDLIKNQCVNFNFNGLTGTGVLTPSSKRFQP
FQQFGRDVSDFTDSVRDPKTSEILDISPCAFGGVSVITPGTNASSEVAVLY
QDVNCTDVSTAIHADQLTPAWRIYSTGNNVFQTQAGCLIGAEHVDTSYE
CDIPIGAGICASYHTVSLLRSTSQKSIVAYTMSLGADSSIAYSNNTIAIPTN
FSISITTEVMPVSMAKTSVDCNMYICGDSTECANLLLQYGSFCTQLNRAL
SGIAAEQDRNTREVFAQVKQMYKTPTLKYFGGFNFSQILPDPLKPTKRSF

IEDLLFNKVTLADAGFMKQYGECLGDINARDLICAQKFNGLTVLPPLLT
DDMIAAYTAALVSGTATAGWTFGAGAALQIPFAMQMAYRFNGIGVTQ
NVLYENQKQIANQFNKAISQIQESLTTTSTALGKLQDVVNQNAQALNTL
VKQLSSNFGAISSVLNDILSRLDKVEAEVQIDRLITGRLQSLQTYVTQQLI
RAAEIRASANLAATKMSECVLGQSKRVDFCGKGYHLMSFPQAAPHGVV
FLHVTYVPSQERNFTTAPAICHEGKAYFPREGVFVFNGTSWFITQRNFFS
PQIITTDNTFVSGNCDVVIGIINNTVYDPLQPELDSFKEELDKYFKNHTSP
DVDLGDISGINASVVNIQKEIDRLNEVAKNLNESLIDLQELGKYEQYIKW
PWYVWLGFIAGLIAIVMVTILLCCMTSCCSCLKGACSCGSCCKFDEDDSE
PVLKGVKLHYT (SEQ ID NO: 1)

IX. <u>Nucleic Acid sequence of a full-length spike (S) protein (nucleotides 1-3768)</u>

ATGTTTATTTTCTTATTATTTCTTACTCTCACTAGTGGTAGTGACCTTG
ACCGGTGCACCACTTTTGATGATGTTCAAGCTCCTAATTACACTCAAC
ATACTTCATCTATGAGGGGGGTTTACTATCCTGATGAAATTTTTAGAT
CAGACACTCTTTATTTAACTCAGGATTTATTTCTTCCATTTTATTCTAA
TGTTACAGGGTTTCATACTATTAATCATACGTTTGGCAACCCTGTCAT
ACCTTTTAAGGATGGTATTTATTTTGCTGCCACAGAGAAATCAAATGT
TGTCCGTGGTTGGGTTTTTGGTTCTACCATGAACAACAAGTCACAGTC
GGTGATTATTATTAACAATTCTACTAATGTTGTTATACGAGCATGTAA
CTTTGAATTGTGTGACAACCCTTTCTTTGCTGTTTCTAAACCCATGGG
TACACAGACACATACTATGATATTCGATAATGCATTTAATTGCACTTT
CGAGTACATATCTGATGCCTTTTCGCTTGATGTTTCAGAAAAGTCAGG
TAATTTTAAACACTTACGAGAGTTTGTGTTTAAAAATAAAGATGGGTT
TCTCTATGTTTATAAGGGCTATCAACCTATAGATGTAGTTCGTGATCT
ACCTTCTGGTTTTAACACTTTGAAACCTATTTTTAAGTTGCCTCTTGGT
ATTAACATTACAAATTTTAGAGCCATTCTTACAGCCTTTTCACCTGCT
CAAGACATTTGGGGCACGTCAGCTGCAGCCTATTTTGTTGGCTATTTA
AAGCCAACTACATTTATGCTCAAGTATGATGAAAATGGTACAATCAC
AGATGCTGTTGATTGTTCTCAAAATCCACTTGCTGAACTCAAATGCTC
TGTTAAGAGCTTTGAGATTGACAAAGGAATTTACCAGACCTCTAATTT
CAGGGTTGTTCCCTCAGGAGATGTTGTGAGATTCCCTAATATTACAAA
CTTGTGTCCTTTTGGAGAGGTTTTTAATGCTACTAAATTCCCTTCTGTC

TATGCATGGGAGAGAAAAAAAATTTCTAATTGTGTTGCTGATTACTCT
GTGCTCTACAACTCAACATTTTTTTCAACCTTTAAGTGCTATGGCGTT
TCTGCCACTAAGTTGAATGATCTTTGCTTCTCCAATGTCTATGCAGAT
TCTTTTGTAGTCAAGGGAGATGATGTAAGACAAATAGCGCCAGGACA
AACTGGTGTTATTGCTGATTATAATTATAAATTGCCAGATGATTTCAT
GGGTTGTGTCCTTGCTTGGAATACTAGGAACATTGATGCTACTTCAAC
TGGTAATTATAATTATAAATATAGGTATCTTAGACATGGCAAGCTTA
GGCCCTTTGAGAGAGACATATCTAATGTGCCTTTCTCCCCTGATGGCA
AACCTTGCACCCCACCTGCTCTTAATTGTTATTGGCCATTAAATGATT
ATGGTTTTTACACCACTACTGGCATTGGCTACCAACCTTACAGAGTTG
TAGTACTTTCTTTTGAACTTTTAAATGCACCGGCCACGGTTTGTGGAC
CAAAATTATCCACTGACCTTATTAAGAACCAGTGTGTCAATTTTAATT
TTAATGGACTCACTGGTACTGGTGTGTTAACTCCTTCTTCAAAGAGAT
TTCAACCATTTCAACAATTTGGCCGTGATGTTTCTGATTTCACTGATT
CCGTTCGAGATCCTAAAACATCTGAAATATTAGACATTTCACCTTGCG
CTTTTGGGGGTGTAAGTGTAATTACACCTGGAACAAATGCTTCATCTG
AAGTTGCTGTTCTATATCAAGATGTTAACTGCACTGATGTTTCTACAG
CAATTCATGCAGATCAACTCACACCAGCTTGGCGCATATATTCTACTG
GAAACAATGTATTCCAGACTCAAGCAGGCTGTCTTATAGGAGCTGAG
CATGTCGACACTTCTTATGAGTGCGACATTCCTATTGGAGCTGGCATT
TGTGCTAGTTACCATACAGTTTCTTTATTACGTAGTACTAGCCAAAAA
TCTATTGTGGCTTATACTATGTCTTTAGGTGCTGATAGTTCAATTGCTT
ACTCTAATAACACCATTGCTATACCTACTAACTTTTCAATTAGCATTA
CTACAGAAGTAATGCCTGTTTCTATGGCTAAAACCTCCGTAGATTGTA
ATATGTACATCTGCGGAGATTCTACTGAATGTGCTAATTTGCTTCTCC
AATATGGTAGCTTTTGCACACAACTAAATCGTGCACTCTCAGGTATTG
CTGCTGAACAGGATCGCAACACACGTGAAGTGTTCGCTCAAGTCAAA
CAAATGTACAAAACCCCAACTTTGAAATATTTTGGTGGTTTTAATTTT
TCACAAATATTACCTGACCCTCTAAAGCCAACTAAGAGGTCTTTTATT
GAGGACTTGCTCTTTAATAAGGTGACACTCGCTGATGCTGGCTTCATG
AAGCAATATGGCGAATGCCTAGGTGATATTAATGCTAGAGATCTCAT
TTGTGCGCAGAAGTTCAATGGACTTACAGTGTTGCCACCTCTGCTCAC
TGATGATATGATTGCTGCCTACACTGCTGCTCTAGTTAGTGGTACTGC
CACTGCTGGATGGACATTTGGTGCTGGCGCTGCTCTTCAAATACCTTT

TGCTATGCAAATGGCATATAGGTTCAATGGCATTGGAGTTACCCAAA
ATGTTCTCTATGAGAACCAAAAACAAATCGCCAACCAATTTAACAAG
GCGATTAGTCAAATTCAAGAATCACTTACAACAACATCAACTGCATT
GGGCAAGCTGCAAGACGTTGTTAACCAGAATGCTCAAGCATTAAACA
CACTTGTTAAACAACTTAGCTCTAATTTTGGTGCAATTTCAAGTGTGC
TAAATGATATCCTTTCGCGACTTGATAAAGTCGAGGCGGAGGTACAA
ATTGACAGGTTAATTACAGGCAGACTTCAAAGCCTTCAAACCTATGT
AACACAACAACTAATCAGGGCTGCTGAAATCAGGGCTTCTGCTAATC
TTGCTGCTACTAAAATGTCTGAGTGTGTTCTTGGACAATCAAAAAGA
GTTGACTTTTGTGGAAAGGGCTACCACCTTATGTCCTTCCCACAAGCA
GCCCCGCATGGTGTTGTCTTCCTACATGTCACGTATGTGCCATCCCAG
GAGAGGAACTTCACCACAGCGCCAGCAATTTGTCATGAAGGCAAAGC
ATACTTCCCTCGTGAAGGTGTTTTTGTGTTTAATGGCACTTCTTGGTTT
ATTACACAGAGGAACTTCTTTTCTCCACAAATAATTACTACAGACAAT
ACATTTGTCTCAGGAAATTGTGATGTCGTTATTGGCATCATTAACAAC
ACAGTTTATGATCCTCTGCAACCTGAGCTCGACTCATTCAAAGAAGA
GCTGGACAAGTACTTCAAAAATCATACATCACCAGATGTTGATCTTG
GCGACATTTCAGGCATTAACGCTTCTGTCGTCAACATTCAAAAAGAA
ATTGACCGCCTCAATGAGGTCGCTAAAAATTTAAATGAATCACTCAT
TGACCTTCAAGAATTGGGAAAATATGAGCAATATATTAAATGGCCTT
GGTATGTTTGGCTCGGCTTCATTGCTGGACTAATTGCCATCGTCATGG
TTACAATCTTGCTTTGTTGCATGACTAGTTGTTGCAGTTGCCTCAAGG
GTGCATGCTCTTGTGGTTCTTGCTGCAAGTTTGATGAGGATGACTCTG
AGCCAGTTCTCAAGGGTGTCAAATTACATTACACATAA (SEQ ID NO: 2)

Example 1

Cloning of the spike protein

The nucleic acid sequence encoding the full length spike protein was obtained through use of overlapping polymerase chain reaction (PCR). Overlapping clones containing fragments of the spike protein were obtained from the British Columbia Cancer Agency (Vancouver, British Columbia). The following primers were used during the PCR reactions to amplify the nucleic acid sequence encoding the full-length spike protein of SARS-CoV: Clone 1:

Forward primer: 5'- A GTC GGA TCC GGT AGG CTT ATC ATT AGA G - 3' (SEQ ID NO: 3); Reverse primer: 5'- CCA TCA GGG GAG AAA GGC AC-3 (SEQ ID NO: 4). Clone 2: Forward primer: 5'- GTG CCT TTC TCC CCT GAT GG-3' (SEQ ID NO: 5); Reverse primer: 5'- GAA GAG CAG CGC CAG CAC C-3' (SEQ ID NO: 6). Clone 3: Forward primer: 5'- GGT GCT GGC GCT GCT CTT C-3' (SEQ ID NO: 7); Reverse primer: 5'- A CTG TCT AGA GTT CGT TTA TGT GTA ATG-3 (SEQ ID NO: 8).

The nucleic acid segment that resulted from overlapping PCR between the nucleic acid segments generated with the above pairs of primers contain amino acid residues from number 1 to number 1255 of the spike protein of the virus (SARS-CoV) that is etiologically linked to severe acute respiratory syndrome. The underlined primer sequences represent restriction enzyme cutting sites for BamHI and XbaI that were used to clone the amplified fragment into pCDNA3(+) (Invitrogen, Carlsbad, California).

The full length spike protein gene has been cloned as shown in Fig. 1. Fig. 1 shows a gel for the nucleic acid segment encoding the full length spike protein inserted into the pCDNA3.1(+) vector that has been digested with the restriction enzymes (Lane 2: BamHI and XbaI; Lane 3: HindIII).

Example 2

Generation of amino-terminal (S1) and carboxyl-terminal (S2) fragments of the full length spike protein

Computer analysis identified a potential functional separation site between the amino-terminus (S1) and the carboxyl-terminus (S2) of the spike protein. The separation site between S1 and S2 is between positions between 758 and 761 (^{758}RNTR^{761}) relative to SEQ ID NO: 1. PCR was used to create nucleic acids that code for the amino-terminal fragment (S1), and the carboxyl-terminal fragment (S2) of the spike protein.

The following primers, S1 forward primer: 5'-AGTC GGA TCC GAC CGG TGC ACC ACT TTT G-3' (SEQ ID NO: 9), and the reverse primer, S1 Reverse primer: 5'-AGTC GGG CCC CTG TTC AGC AGC AAT ACC-3' (SEQ ID NO: 10), were used to prepare a nucleic acid segment coding for amino acid residues 17-757 of the spike protein. Two restriction sites, BamHI and ApaI, underlined in the two primers were used to clone the nucleic acid segment

coding for the amino-terminal fragment of the spike protein (S1) gene into the pSecTag2B plasmid for expression.

The following pair of primers, S2 Forward: 5'-ACTG GGATCC GAA GTG TTC GCT CAA GTC-3' (SEQ ID NO: 11), and S2 Reverse: 5'-ACTG TCTAGA TTG CTC ATA TTT TCC C-3' (SEQ ID NO: 12), were used within a PCR reaction to prepare a nucleic acid segment coding for amino acid residues 762-1189 of the spike protein. Two restriction sites, BamHI and XbaI, underlined in the two primers were used to clone the nucleic acid segment coding for the carboxyl-terminal fragment of the spike protein (S2) gene into pCDNA3.1(+) plasmid for expression.

To create a fragment containing residues 272-537, the following pair of primers was used for PCR amplification: primer 5' GATCGGATCCGGTACAATCACAG 3' (SEQ ID NO:64) and primer 5' GATCGGGCCCGACACACTGGTTC 3' (SEQ ID NO:65). The amplified fragment was digested with *Bam*HI and *Apa*I and ligated into pSecTag2B digested with the same restriction enzymes. A schematic diagram of the position of many of the soluble spike protein fragments within the full-length spike protein is provided in Fig. 1B.

In some cases, nucleic acids encoding the S fragments and full-length S polypeptides had their native leader sequence (spike protein amino acids 1-16, MFIFLLFLTLTSGSDL (SEQ ID NO:60)) replaced with a mouse k chain leader sequence (METDTLLLWVLLLWVPGSTGD) (SEQ ID NO: 16) to permit secretion, as described below.

Example 3

Generation of the whole soluble spike protein (sS) lacking the cytoplasmic tail and the transmembrane domain

The following pair of primers were used to generate a nucleic acid segment encoding a fragment of the spike protein (sS) lacking the cytoplasmic tail having amino acids 17-1189 of SEQ ID NO: 1: S1 Forward: 5'- AGTC GGATCC GAC CGG TGC ACC ACT TTT G-3' (SEQ ID NO: 9), and Reverse: 5' ACTG TCTAGA TTG CTC ATA TTT TCC C-3' (SEQ ID NO: 12).

Example 4

Expression of an amino-terminal and carboxyl-terminal fragment of a spike protein

Expression will be done by transfecting an expression construct containing the pSecTag2B or pCDNA3.1(+) plasmid and a nucleic acid insert that encodes an amino-terminal (S1), a carboxyl-terminal (S2) fragment, or a fragment of the spike protein of SARS-CoV that lacks the cytoplasmic tail and the transmembrane domain, into 293 or Vero E6 cells. It is thought that elimination of the transmembrane domain allows the polypeptides and peptide fragments to be soluble in an aqueous solution. Expression efficiency of the encoded fragments will then be tested. Once a positive signal is obtained as determined with gel analysis, a stably transfected cell line will be generated. The full length spike protein, and fragments thereof will be purified according to methods that are routinely used with other highly glycosylated proteins. Such as use of a lentil lectin column for large production. The resulting proteins: soluble S1 (sS1), soluble S2 (sS2) and whole soluble S (sS) will have the following amino acid sequences. Bold lettering denotes the signal peptide which can be cleaved so the excreted protein will not contain it.

Amino acid sequence of a soluble amino-terminal fragment of the spike protein (amino acids 17-757)

DRCTTFDDVQAPNYTQHTSSMRGVYYPDEIFRSDTLYLTQDLFLPFYSN
VTGFHTINHTFGNPVIPFKDGIYFAATEKSNVVRGWVFGSTMNNKSQSVI
IINNSTNVVIRACNFELCDNPFFAVSKPMGTQTHTMIFDNAFNCTFEYISD
AFSLDVSEKSGNFKHLREFVFKNKDGFLYVYKGYQPIDVVRDLPSGFNT
LKPIFKLPLGINITNFRAILTAFSPAQDIWGTSAAAYFVGYLKPTTFMLKY
DENGTITDAVDCSQNPLAELKCSVKSFEIDKGIYQTSNFRVVPSGDVVRF
PNITNLCPFGEVFNATKFPSVYAWERKKISNCVADYSVLYNSTFFSTFKC
YGVSATKLNDLCFSNVYADSFVVKGDDVRQIAPGQTGVIADYNYKLPD
DFMGCVLAWNTRNIDATSTGNYNYKYRYLRHGKLRPFERDISNVPFSPD
GKPCTPPALNCYWPLNDYGFYTTTGIGYQPYRVVVLSFELLNAPATVCG
PKLSTDLIKNQCVNFNFNGLTGTGVLTPSSKRFQPFQQFGRDVSDFTDSV
RDPKTSEILDISPCAFGGVSVITPGTNASSEVAVLYQDVNCTDVSTAIHAD
QLTPAWRIYSTGNNVFQTQAGCLIGAEHVDTSYECDIPIGAGICASYHTV

SLLRSTSQKSIVAYTMSLGADSSIAYSNNTIAIPTNFSISITTEVMPVSMAK TSVDCNMYICGDSTECANLLLQYGSFCTQLNRALSGIAAEQ (SEQ ID NO: 13)

Amino acid sequence of a soluble carboxyl-terminal fragment of the spike protein (amino acids 762-1189)

EVFAQVKQMYKTPTLKYFGGFNFSQILPDPLKPTKRSFIEDLLFNKVTLA DAGFMKQYGECLGDINARDLICAQKFNGLTVLPPLLTDDMIAAYTAALV SGTATAGWTFGAGAALQIPFAMQMAYRFNGIGVTQNVLYENQKQIANQ FNKAISQIQESLTTTSTALGKLQDVVNQNAQALNTLVKQLSSNFGAISSV LNDILSRLDKVEAEVQIDRLITGRLQSLQTYVTQQLIRAAEIRASANLAAT KMSECVLGQSKRVDFCGKGYHLMSFPQAAPHGVVFLHVTYVPSQERNF TTAPAICHEGKAYFPREGVFVFNGTSWFITQRNFFSPQIITTDNTFVSGNC DVVIGIINNTVYDPLQPELDSFKEELDKYFKNHTSPDVDLGDISGINASVV NIQKEIDRLNEVAKNLNESLIDLQELGKYEQ (SEQ ID NO: 14)

Amino acid sequence of a soluble spike protein having amino acids 17-757 and 762-1189 of SEQ ID NO: 1 (lacking the signal peptide and the potential cleavage site)

DRCTTFDDVQAPNYTQHTSSMRGVYYPDEIFRSDTLYLTQDLFLPFYSN VTGFHTINHTFGNPVIPFKDGIYFAATEKSNVVRGWVFGSTMNNKSQSVI IINNSTNVVIRACNFELCDNPFFAVSKPMGTQTHTMIFDNAFNCTFEYISD AFSLDVSEKSGNFKHLREFVFKNKDGFLYVYKGYQPIDVVRDLPSGFNT LKPIFKLPLGINITNFRAILTAFSPAQDIWGTSAAAYFVGYLKPTTFMLKY DENGTITDAVDCSQNPLAELKCSVKSFEIDKGIYQTSNFRVVPSGDVVRF PNITNLCPFGEVFNATKFPSVYAWERKKISNCVADYSVLYNSTFFSTFKC YGVSATKLNDLCFSNVYADSFVVKGDDVRQIAPGQTGVIADYNYKLPD DFMGCVLAWNTRNIDATSTGNYNYKYRYLRHGKLRPFERDISNVPFSPD GKPCTPPALNCYWPLNDYGFYTTTGIGYQPYRVVVLSFELLNAPATVCG PKLSTDLIKNQCVNFNFNGLTGTGVLTPSSKRFQPFQQFGRDVSDFTDSV RDPKTSEILDISPCAFGGVSVITPGTNASSEVAVLYQDVNCTDVSTAIHAD QLTPAWRIYSTGNNVFQTQAGCLIGAEHVDTSYECDIPIGAGICASYHTV SLLRSTSQKSIVAYTMSLGADSSIAYSNNTIAIPTNFSISITTEVMPVSMAK TSVDCNMYICGDSTECANLLLQYGSFCTQLNRALSGIAAEQDEVFAQVK

QMYKTPTLKYFGGFNFSQILPDPLKPTKRSFIEDLLFNKVTLADAGFMKQ YGECLGDINARDLICAQKFNGLTVLPPLLTDDMIAAYTAALVSGTATAG WTFGAGAALQIPFAMQMAYRFNGIGVTQNVLYENQKQIANQFNKAISQI QESLTTTSTALGKLQDVVNQNAQALNTLVKQLSSNFGAISSVLNDILSRL DKVEAEVQIDRLITGRLQSLQTYVTQQLIRAAEIRASANLAATKMSECVL GQSKRVDFCGKGYHLMSFPQAAPHGVVFLHVTYVPSQERNFTTAPAICH EGKAYFPREGVFVFNGTSWFITQRNFFSPQIITTDNTFVSGNCDVVIGIINN TVYDPLQPELDSFKEELDKYFKNHTSPDVDLGDISGINASVVNIQKEIDRL NEVAKNLNESLIDLQELGKYEQ (SEQ ID NO: 15)

Example 5

Generation of additional soluble fragments of the spike protein

The nucleic acid sequence encoding a polypeptide containing amino acids 17-757 of SEQ ID NO: 1 was obtained through use of polymerase chain reaction (PCR). The following primers were used during the PCR reactions to amplify the nucleic acid sequence: Forward primer: 5' AGCT GGA TCC GAC CGG TGC ACC ACT TTT G 3' (SEQ ID NO: 9); and Reverse primer: 5' AGCT GGG CCC CTG TTC AGC AGC AAT ACC 3' (SEQ ID NO: 10). The resulting PCR product was digested with BamHI and ApaI and, encodes a polypeptide having an amino acid sequence corresponding to SEQ ID NO: 43. The digested PCR product was then ligated to pSecTag2B (Invitrogen, Carlsbad, California) that was digested with the same enzymes. The pSecTag2B construct containing the PCR product insert encodes a polypeptide having SEQ ID NO: 46 with the mouse k chain leader sequence (METDTLLLWVLLLWVPGSTGD) (SEQ ID NO: 16) at the N-terminus for secretion, and a myc epitope (EQKLISEEDL) (SEQ ID NO: 17) plus a histidine tag (HHHHHH) (SEQ ID NO: 18) at the C-terminus for affinity purification.

The nucleic acid sequence encoding a polypeptide containing amino acids 17-276 of SEQ ID NO: 1 was obtained through use of polymerase chain reaction (PCR). The following primers were used during the PCR reactions to amplify the nucleic acid sequence: Forward primer: 5' AGCT GGA TCC GAC CGG TGC ACC ACT TTT G 3' (SEQ ID NO: 9); and Reverse primer: 5' CTAG CTC GAG CAA CAG CAT CTG TG 3' (SEQ ID NO: 19). The resulting PCR product was digested with BamHI and XhoI and, encodes an

amino acid having SEQ ID NO: 44. The digested PCR product was then ligated to pSecTag2B (Invitrogen, Carlsbad, California) that was digested with the same enzymes. The pSecTag2B construct containing the PCR product insert encodes a polypeptide having SEQ ID NO: 47 with the mouse k chain leader sequence (METDTLLLWVLLLWVPGSTGD) (SEQ ID NO: 16) at the N-terminus for secretion, and a myc epitope (EQKLISEEDL) (SEQ ID NO: 17) plus a histidine tag (HHHHHH) (SEQ ID NO: 18) at the C-terminus for affinity purification.

The nucleic acid sequence encoding a polypeptide containing amino acids 17-537 of SEQ ID NO: 1 was obtained by digesting the nucleic acid sequence that encodes SEQ ID NO: 43 (as described above) with BamHI and HincII. The nucleic acid segment produced encodes a polypeptide having SEQ ID NO: 45. This nucleic acid segment was ligated into a pSecTag2B vector that was digested with BamHI and EcoRV. The pSecTag2B construct containing the PCR product insert encodes a polypeptide having SEQ ID NO: 48 with the mouse k chain leader sequence (METDTLLLWVLLLWVPGSTGD) (SEQ ID NO: 16) at the N-terminus for secretion, and a myc epitope (EQKLISEEDL) (SEQ ID NO: 17) plus a histidine tag (HHHHHH) (SEQ ID NO: 18) at the C-terminus for affinity purification.

The expression of these peptide fragments in mammalian cells is illustrated in Fig. 3. This figure shows that the peptide fragments can be secreted into medium in which cells that express the peptide fragments are grown. Fig. 3 also indicates that the peptide fragments are soluble in aqueous medium.

Table 1

Examples of additional peptide fragments of the invention

SEQ ID NUMBER	Amino Acid Position Relative to SEQ ID NO: 1	Amino Acid Sequence Amino acid residues in bold (amino acid residues 1-16) identify a signal sequence.
20	1-100	**MFIFLLFLTLTSGSDL**DRCTTFDDVQAPNYTQHTSSMRGVYYPDEIFRSDTLYLTQDLFLPFYSNVTGFHTINHTFGNPVIPFKDGIYFAATEKSNVVRG
21	101-200	WVFGSTMNNKSQSVIIINNSTNVVIRACNFELCDNPFFAVSKPMGTQTHTMIFDNAFNCTFEYISDAFSLDVSEKSGNFKHLREFVFKNKDGFLYVYKGY
22	201-300	QPIDVVRDLPSGFNTLKPIFKLPLGINITNFRAILTAFSPAQDIWGTSAAAYFVGYLKPTTFMLKYDENGTITDAVDCSQNPLAELKCSVKSFEIDKGIY
23	301-400	QTSNFRVVPSGDVVRFPNITNLCPFGEVFNATKFPSVYAWERKKISNCVADYSVLYNSTFFSTFKCYGVSATKLNDLCFSNVYADSFVVKGDDVRQIAPG
24	401-500	QTGVIADYNYKLPDDFMGCVLAWNTRNIDATSTGNYNYKYRYLRHGKLRPFERDISNVPFSPDGKPCTPPALNCYWPLNDYGFYTTTGIGYQPYRVVVLS
25	501-600	FELLNAPATVCGPKLSTDLIKNQCVNFNFNGLTGTGVLTPSSKRFQPFQQFGRDVSDFTDSVRDPKTSEILDISPCAFGGVSVITPGTNASSEVAVLYQD
26	601-700	VNCTDVSTAIHADQLTPAWRIYSTGNNVFQTQAGCLIGAEHVDTSYECDIPIGAGICASYHTVSLLRSTSQKSIVAYTMSLGADSSIAYSNNTIAIPTNF
27	701-800	SISITTEVMPVSMAKTSVDCNMYICGDSTECANLLLQYGSFCTQLNRALSGIAAEQDRNTREVFAQVKQMYKTPTLKYFGGFNFSQILPDPLKPTKRSFI
28	801-900	EDLLFNKVTLADAGFMKQYGECLGDINARDLICAQKFNGLTVLPPLLTDDMIAAYTAALVSGTATAGWTFGAGAALQIPFAMQMAYRFNGIGVTQNVLYE

SEQ ID NUMBER	Amino Acid Position Relative to SEQ ID NO: 1	Amino Acid Sequence Amino acid residues in bold (amino acid residues 1-16) identify a signal sequence.
29	901-1000	NQKQIANQFNKAISQIQESLTTTSTALGK LQDVVNQNAQALNTLVKQLSSNFGAISS VLNDILSRLDKVEAEVQIDRLITGRLQSL QTYVTQQLIRAAEI
30	1001-1100	RASANLAATKMSECVLGQSKRVDFCGK GYHLMSFPQAAPHGVVFLHVTYVPSQE RNFTTAPAICHEGKAYFPREGVFVFNGT SWFITQRNFFSPQIITTD
31	1101-1189	NTFVSGNCDVVIGIINNTVYDPLQPELDS FKEELDKYFKNHTSPDVDLGDISGINASV VNIQKEIDRLNEVAKNLNESLIDLQELGK YEQ
32	1-200	**MFIFLLFLTLTSGSDL**DRCTTFDDVQAP NYTQHTSSMRGVYYPDEIFRSDTLYLTQ DLFLPFYSNVTGFHTINHTFGNPVIPFKD GIYFAATEKSNVVRGWVFGSTMNNKSQ SVIIINNSTNVVIRACNFELCDNPFFAVSK PMGTQTHTMIFDNAFNCTFEYISDAFSLD VSEKSGNFKHLREFVFKNKDGFLYVYK GY
33	201-400	QPIDVVRDLPSGFNTLKPIFKLPLGINITN FRAILTAFSPAQDIWGTSAAAYFVGYLK PTTFMLKYDENGTITDAVDCSQNPLAEL KCSVKSFEIDKGIYQTSNFRVVPSGDVVR FPNITNLCPFGEVFNATKFPSVYAWERK KISNCVADYSVLYNSTFFSTFKCYGVSA TKLNDLCFSNVYADSFVVKGDDVRQIAP G
34	401-600	QTGVIADYNYKLPDDFMGCVLAWNTRN IDATSTGNYNYKYRYLRHGKLRPFERDI SNVPFSPDGKPCTPPALNCYWPLNDYGF YTTTGIGYQPYRVVVLSFELLNAPATVC GPKLSTDLIKNQCVNFNFNGLTGTGVLT PSSKRFQPFQQFGRDVSDFTDSVRDPKTS EILDISPCAFGGVSVITPGTNASSEVAVLY QD

SEQ ID NUMBER	Amino Acid Position Relative to SEQ ID NO: 1	Amino Acid Sequence Amino acid residues in bold (amino acid residues 1-16) identify a signal sequence.
35	601-800	VNCTDVSTAIHADQLTPAWRIYSTGNNV FQTQAGCLIGAEHVDTSYECDIPIGAGIC ASYHTVSLLRSTSQKSIVAYTMSLGADS SIAYSNNTIAIPTNFSISITTEVMPVSMAK TSVDCNMYICGDSTECANLLLQYGSFCT QLNRALSGIAAEQDRNTREVFAQVKQM YKTPTLKYFGGFNFSQILPDPLKPTKRSFI
36	801-1000	EDLLFNKVTLADAGFMKQYGECLGDIN ARDLICAQKFNGLTVLPPLLTDDMIAAY TAALVSGTATAGWTFGAGAALQIPFAM QMAYRFNGIGVTQNVLYENQKQIANQF NKAISQIQESLTTTSTALGKLQDVVNQN AQALNTLVKQLSSNFGAISSVLNDILSRL DKVEAEVQIDRLITGRLQSLQTYVTQQLI RAAEI
37	1001-1189	RASANLAATKMSECVLGQSKRVDFCGK GYHLMSFPQAAPHGVVFLHVTYVPSQE RNFTTAPAICHEGKAYFPREGVFVFNGT SWFITQRNFFSPQIITTDNTFVSGNCDVVI GIINNTVYDPLQPELDSFKEELDKYFKNH TSPDVDLGDISGINASVVNIQKEIDRLNE VAKNLNESLIDLQELGKYEQ
38	1-400	**MFIFLLFLTLTSGSDL**DRCTTFDDVQAP NYTQHTSSMRGVYYPDEIFRSDTLYLTQ DLFLPFYSNVTGFHTINHTFGNPVIPFKD GIYFAATEKSNVVRGWVFGSTMNNKSQ SVIIINNSTNVVIRACNFELCDNPFFAVSK PMGTQTHTMIFDNAFNCTFEYISDAFSLD VSEKSGNFKHLREFVFKNKDGFLYVYK GYQPIDVVRDLPSGFNTLKPIFKLPLGINI TNFRAILTAFSPAQDIWGTSAAAYFVGY LKPTTFMLKYDENGTITDAVDCSQNPLA ELKCSVKSFEIDKGIYQTSNFRVVPSGDV VRFPNITNLCPFGEVFNATKFPSVYAWE RKKISNCVADYSVLYNSTFFSTFKCYGV SATKLNDLCFSNVYADSFVVKGDDVRQI APG

SEQ ID NUMBER	Amino Acid Position Relative to SEQ ID NO: 1	Amino Acid Sequence Amino acid residues in bold (amino acid residues 1-16) identify a signal sequence.
39	1-600	**MFIFLLFLTLTSGSDL**DRCTTFDDVQAP NYTQHTSSMRGVYYPDEIFRSDTLYLTQ DLFLPFYSNVTGFHTINHTFGNPVIPFKD GIYFAATEKSNVVRGWVFGSTMNNKSQ SVIIINNSTNVVIRACNFELCDNPFFAVSK PMGTQTHTMIFDNAFNCTFEYISDAFSLD VSEKSGNFKHLREFVFKNKDGFLYVYK GYQPIDVVRDLPSGFNTLKPIFKLPLGINI TNFRAILTAFSPAQDIWGTSAAAYFVGY LKPTTFMLKYDENGTITDAVDCSQNPLA ELKCSVKSFEIDKGIYQTSNFRVVPSGDV VRFPNITNLCPFGEVFNATKFPSVYAWE RKKISNCVADYSVLYNSTFFSTFKCYGV SATKLNDLCFSNVYADSFVVKGDDVRQI APGQTGVIADYNYKLPDDFMGCVLAWN TRNIDATSTGNYNYKYRYLRHGKLRPFE RDISNVPFSPDGKPCTPPALNCYWPLND YGFYTTTGIGYQPYRVVVLSFELLNAPA TVCGPKLSTDLIKNQCVNFNFNGLTGTG VLTPSSKRFQPFQQFGRDVSDFTDSVRDP KTSEILDISPCAFGGVSVITPGTNASSEVA VLYQD

SEQ ID NUMBER	Amino Acid Position Relative to SEQ ID NO: 1	Amino Acid Sequence Amino acid residues in bold (amino acid residues 1-16) identify a signal sequence.
40	1-800	**MFIFLLFLTLTSGSDL**DRCTTFDDVQAP NYTQHTSSMRGVYYPDEIFRSDTLYLTQ DLFLPFYSNVTGFHTINHTFGNPVIPFKD GIYFAATEKSNVVRGWVFGSTMNNKSQ SVIIINNSTNVVIRACNFELCDNPFFAVSK PMGTQTHTMIFDNAFNCTFEYISDAFSLD VSEKSGNFKHLREFVFKNKDGFLYVYK GYQPIDVVRDLPSGFNTLKPIFKLPLGINI TNFRAILTAFSPAQDIWGTSAAAYFVGY LKPTTFMLKYDENGTITDAVDCSQNPLA ELKCSVKSFEIDKGIYQTSNFRVVPSGDV VRFPNITNLCPFGEVFNATKFPSVYAWE RKKISNCVADYSVLYNSTFFSTFKCYGV SATKLNDLCFSNVYADSFVVKGDDVRQI APGQTGVIADYNYKLPDDFMGCVLAWN TRNIDATSTGNYNYKYRYLRHGKLRPFE RDISNVPFSPDGKPCTPPALNCYWPLND YGFYTTTGIGYQPYRVVVLSFELLNAPA TVCGPKLSTDLIKNQCVNFNFNGLTGTG VLTPSSKRFQPFQQFGRDVSDFTDSVRDP KTSEILDISPCAFGGVSVITPGTNASSEVA VLYQDVNCTDVSTAIHADQLTPAWRIYS TGNNVFQTQAGCLIGAEHVDTSYECDIPI GAGICASYHTVSLLRSTSQKSIVAYTMSL GADSSIAYSNNTIAIPTNFSISITTEVMPVS MAKTSVDCNMYICGDSTECANLLLQYG SFCTQLNRALSGIAAEQDRNTREVFAQV KQMYKTPTLKYFGGFNFSQILPDPLKPT KRSFI

SEQ ID NUMBER	Amino Acid Position Relative to SEQ ID NO: 1	Amino Acid Sequence Amino acid residues in bold (amino acid residues 1-16) identify a signal sequence.
41	1-1000	**MFIFLLFLTLTSGSDL**DRCTTFDDVQAP NYTQHTSSMRGVYYPDEIFRSDTLYLTQ DLFLPFYSNVTGFHTINHTFGNPVIPFKD GIYFAATEKSNVVRGWVFGSTMNNKSQ SVIIINNSTNVVIRACNFELCDNPFFAVSK PMGTQTHTMIFDNAFNCTFEYISDAFSLD VSEKSGNFKHLREFVFKNKDGFLYVYK GYQPIDVVRDLPSGFNTLKPIFKLPLGINI TNFRAILTAFSPAQDIWGTSAAAYFVGY LKPTTFMLKYDENGTITDAVDCSQNPLA ELKCSVKSFEIDKGIYQTSNFRVVPSGDV VRFPNITNLCPFGEVFNATKFPSVYAWE RKKISNCVADYSVLYNSTFFSTFKCYGV SATKLNDLCFSNVYADSFVVKGDDVRQI APGQTGVIADYNYKLPDDFMGCVLAWN TRNIDATSTGNYNYKYRYLRHGKLRPFE RDISNVPFSPDGKPCTPPALNCYWPLND YGFYTTTGIGYQPYRVVVLSFELLNAPA TVCGPKLSTDLIKNQCVNFNFNGLTGTG VLTPSSKRFQPFQQFGRDVSDFTDSVRDP KTSEILDISPCAFGGVSVITPGTNASSEVA VLYQDVNCTDVSTAIHADQLTPAWRIYS TGNNVFQTQAGCLIGAEHVDTSYECDIPI GAGICASYHTVSLLRSTSQKSIVAYTMSL GADSSIAYSNNTIAIPTNFSISITTEVMPVS MAKTSVDCNMYICGDSTECANLLLQYG SFCTQLNRALSGIAAEQDRNTREVFAQV KQMYKTPTLKYFGGFNFSQILPDPLKPT KRSFIEDLLFNKVTLADAGFMKQYGECL GDINARDLICAQKFNGLTVLPPLLTDDMI AAYTAALVSGTATAGWTFGAGAALQIP FAMQMAYRFNGIGVTQNVLYENQKQIA NQFNKAISQIQESLTTTSTALGKLQDVVN QNAQALNTLVKQLSSNFGAISSVLNDILS RLDKVEAEVQIDRLITGRLQSLQTYVTQ QLIRAAEI

SEQ ID NUMBER	Amino Acid Position Relative to SEQ ID NO: 1	Amino Acid Sequence Amino acid residues in bold (amino acid residues 1-16) identify a signal sequence.
42	1-1189	**MFIFLLFLTLTSGSDL**DRCTTFDDVQAP NYTQHTSSMRGVYYPDEIFRSDTLYLTQ DLFLPFYSNVTGFHTINHTFGNPVIPFKD GIYFAATEKSNVVRGWVFGSTMNNKSQ SVIIINNSTNVVIRACNFELCDNPFFAVSK PMGTQTHTMIFDNAFNCTFEYISDAFSLD VSEKSGNFKHLREFVFKNKDGFLYVYK GYQPIDVVRDLPSGFNTLKPIFKLPLGINI TNFRAILTAFSPAQDIWGTSAAAYFVGY LKPTTFMLKYDENGTITDAVDCSQNPLA ELKCSVKSFEIDKGIYQTSNFRVVPSGDV VRFPNITNLCPFGEVFNATKFPSVYAWE RKKISNCVADYSVLYNSTFFSTFKCYGV SATKLNDLCFSNVYADSFVVKGDDVRQI APGQTGVIADYNYKLPDDFMGCVLAWN TRNIDATSTGNYNYKYRYLRHGKLRPFE RDISNVPFSPDGKPCTPPALNCYWPLND YGFYTTTGIGYQPYRVVVLSFELLNAPA TVCGPKLSTDLIKNQCVNFNFNGLTGTG VLTPSSKRFQPFQQFGRDVSDFTDSVRDP KTSEILDISPCAFGGVSVITPGTNASSEVA VLYQDVNCTDVSTAIHADQLTPAWRIYS TGNNVFQTQAGCLIGAEHVDTSYECDIPI GAGICASYHTVSLLRSTSQKSIVAYTMSL GADSSIAYSNNTIAIPTNFSISITTEVMPVS MAKTSVDCNMYICGDSTECANLLLQYG SFCTQLNRALSGIAAEQDRNTREVFAQV KQMYKTPTLKYFGGFNFSQILPDPLKPT KRSFIEDLLFNKVTLADAGFMKQYGECL GDINARDLICAQKFNGLTVLPPLLTDDMI AAYTAALVSGTATAGWTFGAGAALQIP FAMQMAYRFNGIGVTQNVLYENQKQIA NQFNKAISQIQESLTTTSTALGKLQDVVN QNAQALNTLVKQLSSNFGAISSVLNDILS RLDKVEAEVQIDRLITGRLQSLQTYVTQ QLIRAAEIRASANLAATKMSECVLGQSK RVDFCGKGYHLMSFPQAAPHGVVFLHV TYVPSQERNFTTAPAICHEGKAYFPREG VFVFNGTSWFITQRNFFSPQIITTDNTFVS GNCDVVIGIINNTVYDPLQPELDSFKEEL DKYFKNHTSPDVDLGDISGINASVVNIQ KEIDRLNEVAKNLNESLIDLQELGKYEQ
43	17-100	DRCTTFDDVQAPNYTQHTSSMRGVYYP DEIFRSDTLYLTQDLFLPFYSNVTGFHTI NHTFGNPVIPFKDGIYFAATEKSNVVRG

SEQ ID NUMBER	Amino Acid Position Relative to SEQ ID NO: 1	Amino Acid Sequence Amino acid residues in bold (amino acid residues 1-16) identify a signal sequence.
44	17-200	DRCTTFDDVQAPNYTQHTSSMRGVYYP DEIFRSDTLYLTQDLFLPFYSNVTGFHTI NHTFGNPVIPFKDGIYFAATEKSNVVRG WVFGSTMNNKSQSVIIINNSTNVVIRACN FELCDNPFFAVSKPMGTQTHTMIFDNAF NCTFEYISDAFSLDVSEKSGNFKHLREFV FKNKDGFLYVYKGY
45	17-400	DRCTTFDDVQAPNYTQHTSSMRGVYYP DEIFRSDTLYLTQDLFLPFYSNVTGFHTI NHTFGNPVIPFKDGIYFAATEKSNVVRG WVFGSTMNNKSQSVIIINNSTNVVIRACN FELCDNPFFAVSKPMGTQTHTMIFDNAF NCTFEYISDAFSLDVSEKSGNFKHLREFV FKNKDGFLYVYKGYQPIDVVRDLPSGFN TLKPIFKLPLGINITNFRAILTAFSPAQDIW GTSAAAYFVGYLKPTTFMLKYDENGTIT DAVDCSQNPLAELKCSVKSFEIDKGIYQ TSNFRVVPSGDVVRFPNITNLCPFGEVFN ATKFPSVYAWERKKISNCVADYSVLYNS TFFSTFKCYGVSATKLNDLCFSNVYADS FVVKGDDVRQIAPG
46	17-600	DRCTTFDDVQAPNYTQHTSSMRGVYYP DEIFRSDTLYLTQDLFLPFYSNVTGFHTI NHTFGNPVIPFKDGIYFAATEKSNVVRG WVFGSTMNNKSQSVIIINNSTNVVIRACN FELCDNPFFAVSKPMGTQTHTMIFDNAF NCTFEYISDAFSLDVSEKSGNFKHLREFV FKNKDGFLYVYKGYQPIDVVRDLPSGFN TLKPIFKLPLGINITNFRAILTAFSPAQDIW GTSAAAYFVGYLKPTTFMLKYDENGTIT DAVDCSQNPLAELKCSVKSFEIDKGIYQ TSNFRVVPSGDVVRFPNITNLCPFGEVFN ATKFPSVYAWERKKISNCVADYSVLYNS TFFSTFKCYGVSATKLNDLCFSNVYADS FVVKGDDVRQIAPGQTGVIADYNYKLPD DFMGCVLAWNTRNIDATSTGNYNYKYR YLRHGKLRPFERDISNVPFSPDGKPCTPP ALNCYWPLNDYGFYTTTGIGYQPYRVV VLSFELLNAPATVCGPKLSTDLIKNQCV NFNFNGLTGTGVLTPSSKRFQPFQQFGR DVSDFTDSVRDPKTSEILDISPCAFGGVS VITPGTNASSEVAVLYQD

SEQ ID NUMBER	Amino Acid Position Relative to SEQ ID NO: 1	Amino Acid Sequence Amino acid residues in bold (amino acid residues 1-16) identify a signal sequence.
47	17-800	DRCTTFDDVQAPNYTQHTSSMRGVYYP DEIFRSDTLYLTQDLFLPFYSNVTGFHTI NHTFGNPVIPFKDGIYFAATEKSNVVRG WVFGSTMNNKSQSVIIINNSTNVVIRACN FELCDNPFFAVSKPMGTQTHTMIFDNAF NCTFEYISDAFSLDVSEKSGNFKHLREFV FKNKDGFLYVYKGYQPIDVVRDLPSGFN TLKPIFKLPLGINITNFRAILTAFSPAQDIW GTSAAAYFVGYLKPTTFMLKYDENGTIT DAVDCSQNPLAELKCSVKSFEIDKGIYQ TSNFRVVPSGDVVRFPNITNLCPFGEVFN ATKFPSVYAWERKKISNCVADYSVLYNS TFFSTFKCYGVSATKLNDLCFSNVYADS FVVKGDDVRQIAPGQTGVIADYNYKLPD DFMGCVLAWNTRNIDATSTGNYNYKYR YLRHGKLRPFERDISNVPFSPDGKPCTPP ALNCYWPLNDYGFYTTTGIGYQPYRVV VLSFELLNAPATVCGPKLSTDLIKNQCV NFNFNGLTGTGVLTPSSKRFQPFQQFGR DVSDFTDSVRDPKTSEILDISPCAFGGVS VITPGTNASSEVAVLYQDVNCTDVSTAI HADQLTPAWRIYSTGNNVFQTQAGCLIG AEHVDTSYECDIPIGAGICASYHTVSLLR STSQKSIVAYTMSLGADSSIAYSNNTIAIP TNFSISITTEVMPVSMAKTSVDCNMYICG DSTECANLLLQYGSFCTQLNRALSGIAA EQDRNTREVFAQVKQMYKTPTLKYFGG FNFSQILPDPLKPTKRSFI

SEQ ID NUMBER	Amino Acid Position Relative to SEQ ID NO: 1	Amino Acid Sequence Amino acid residues in bold (amino acid residues 1-16) identify a signal sequence.
48	17-1000	DRCTTFDDVQAPNYTQHTSSMRGVYYP DEIFRSDTLYLTQDLFLPFYSNVTGFHTI NHTFGNPVIPFKDGIYFAATEKSNVVRG WVFGSTMNNKSQSVIIINNSTNVVIRACN FELCDNPFFAVSKPMGTQTHTMIFDNAF NCTFEYISDAFSLDVSEKSGNFKHLREFV FKNKDGFLYVYKGYQPIDVVRDLPSGFN TLKPIFKLPLGINITNFRAILTAFSPAQDIW GTSAAAYFVGYLKPTTFMLKYDENGTIT DAVDCSQNPLAELKCSVKSFEIDKGIYQ TSNFRVVPSGDVVRFPNITNLCPFGEVFN ATKFPSVYAWERKKISNCVADYSVLYNS TFFSTFKCYGVSATKLNDLCFSNVYADS FVVKGDDVRQIAPGQTGVIADYNYKLPD DFMGCVLAWNTRNIDATSTGNYNYKYR YLRHGKLRPFERDISNVPFSPDGKPCTPP ALNCYWPLNDYGFYTTTGIGYQPYRVV VLSFELLNAPATVCGPKLSTDLIKNQCV NFNFNGLTGTGVLTPSSKRFQPFQQFGR DVSDFTDSVRDPKTSEILDISPCAFGGVS VITPGTNASSEVAVLYQDVNCTDVSTAI HADQLTPAWRIYSTGNNVFQTQAGCLIG AEHVDTSYECDIPIGAGICASYHTVSLLR STSQKSIVAYTMSLGADSSIAYSNNTIAIP TNFSISITTEVMPVSMAKTSVDCNMYICG DSTECANLLLQYGSFCTQLNRALSGIAA EQDRNTREVFAQVKQMYKTPTLKYFGG FNFSQILPDPLKPTKRSFIEDLLFNKVTLA DAGFMKQYGECLGDINARDLICAQKFN GLTVLPPLLTDDMIAAYTAALVSGTATA GWTFGAGAALQIPFAMQMAYRFNGIGV TQNVLYENQKQIANQFNKAISQIQESLTT TSTALGKLQDVVNQNAQALNTLVKQLS SNFGAISSVLNDILSRLDKVEAEVQIDRLI TGRLQSLQTYVTQQLIRAAEI

SEQ ID NUMBER	Amino Acid Position Relative to SEQ ID NO: 1	Amino Acid Sequence Amino acid residues in bold (amino acid residues 1-16) identify a signal sequence.
49	17-1189	DRCTTFDDVQAPNYTQHTSSMRGVYYP DEIFRSDTLYLTQDLFLPFYSNVTGFHTI NHTFGNPVIPFKDGIYFAATEKSNVVRG WVFGSTMNNKSQSVIIINNSTNVVIRACN FELCDNPFFAVSKPMGTQTHTMIFDNAF NCTFEYISDAFSLDVSEKSGNFKHLREFV FKNKDGFLYVYKGYQPIDVVRDLPSGFN TLKPIFKLPLGINITNFRAILTAFSPAQDIW GTSAAAYFVGYLKPTTFMLKYDENGTIT DAVDCSQNPLAELKCSVKSFEIDKGIYQ TSNFRVVPSGDVVRFPNITNLCPFGEVFN ATKFPSVYAWERKKISNCVADYSVLYNS TFFSTFKCYGVSATKLNDLCFSNVYADS FVVKGDDVRQIAPGQTGVIADYNYKLPD DFMGCVLAWNTRNIDATSTGNYNYKYR YLRHGKLRPFERDISNVPFSPDGKPCTPP ALNCYWPLNDYGFYTTTGIGYQPYRVV VLSFELLNAPATVCGPKLSTDLIKNQCV NFNFNGLTGTGVLTPSSKRFQPFQQFGR DVSDFTDSVRDPKTSEILDISPCAFGGVS VITPGTNASSEVAVLYQDVNCTDVSTAI HADQLTPAWRIYSTGNNVFQTQAGCLIG AEHVDTSYECDIPIGAGICASYHTVSLLR STSQKSIVAYTMSLGADSSIAYSNNTIAIP TNFSISITTEVMPVSMAKTSVDCNMYICG DSTECANLLLQYGSFCTQLNRALSGIAA EQDRNTREVFAQVKQMYKTPTLKYFGG FNFSQILPDPLKPTKRSFIEDLLFNKVTLA DAGFMKQYGECLGDINARDLICAQKFN GLTVLPPLLTDDMIAAYTAALVSGTATA GWTFGAGAALQIPFAMQMAYRFNGIGV TQNVLYENQKQIANQFNKAISQIQESLTT TSTALGKLQDVVNQNAQALNTLVKQLS SNFGAISSVLNDILSRLDKVEAEVQIDRLI TGRLQSLQTYVTQQLIRAAEIRASANLA ATKMSECVLGQSKRVDFCGKGYHLMSF PQAAPHGVVFLHVTYVPSQERNFTTAPA ICHEGKAYFPREGVFVFNGTSWFITQRNF FSPQIITTDNTFVSGNCDVVIGIINNTVYD PLQPELDSFKEELDKYFKNHTSPDVDLG DISGINASVVNIQKEIDRLNEVAKNLNES LIDLQELGKYEQ

SEQ ID NUMBER	Amino Acid Position Relative to SEQ ID NO: 1	Amino Acid Sequence Amino acid residues in bold (amino acid residues 1-16) identify a signal sequence.
50	17-276	DRCTTFDDVQAPNYTQHTSSMRGVYYP DEIFRSDTLYLTQDLFLPFYSNVTGFHTI NHTFGNPVIPFKDGIYFAATEKSNVVRG WVFGSTMNNKSQSVIIINNSTNVVIRACN FELCDNPFFAVSKPMGTQTHTMIFDNAF NCTFEYISDAFSLDVSEKSGNFKHLREFV FKNKDGFLYVYKGYQPIDVVRDLPSGFN TLKPIFKLPLGINITNFRAILTAFSPAQDIW GTSAAAYFVGYLKPTTFMLKYDENGTIT DAV
51	17-446	DRCTTFDDVQAPNYTQHTSSMRGVYYP DEIFRSDTLYLTQDLFLPFYSNVTGFHTI NHTFGNPVIPFKDGIYFAATEKSNVVRG WVFGSTMNNKSQSVIIINNSTNVVIRACN FELCDNPFFAVSKPMGTQTHTMIFDNAF NCTFEYISDAFSLDVSEKSGNFKHLREFV FKNKDGFLYVYKGYQPIDVVRDLPSGFN TLKPIFKLPLGINITNFRAILTAFSPAQDIW GTSAAAYFVGYLKPTTFMLKYDENGTIT DAVDCSQNPLAELKCSVKSFEIDKGIYQ TSNFRVVPSGDVVRFPNITNLCPFGEVFN ATKFPSVYAWERKKISNCVADYSVLYNS TFFSTFKCYGVSATKLNDLCFSNVYADS FVVKGDDVRQIAPGQTGVIADYNYKLPD DFMGCVLAWNTRNIDATSTGNYNYKYR YLRHG

SEQ ID NUMBER	Amino Acid Position Relative to SEQ ID NO: 1	Amino Acid Sequence Amino acid residues in bold (amino acid residues 1-16) identify a signal sequence.
52	17-537	DRCTTFDDVQAPNYTQHTSSMRGVYYP DEIFRSDTLYLTQDLFLPFYSNVTGFHTI NHTFGNPVIPFKDGIYFAATEKSNVVRG WVFGSTMNNKSQSVIIINNSTNVVIRACN FELCDNPFFAVSKPMGTQTHTMIFDNAF NCTFEYISDAFSLDVSEKSGNFKHLREFV FKNKDGFLYVYKGYQPIDVVRDLPSGFN TLKPIFKLPLGINITNFRAILTAFSPAQDIW GTSAAAYFVGYLKPTTFMLKYDENGTIT DAVDCSQNPLAELKCSVKSFEIDKGIYQ TSNFRVVPSGDVVRFPNITNLCPFGEVFN ATKFPSVYAWERKKISNCVADYSVLYNS TFFSTFKCYGVSATKLNDLCFSNVYADS FVVKGDDVRQIAPGQTGVIADYNYKLPD DFMGCVLAWNTRNIDATSTGNYNYKYR YLRHGKLRPFERDISNVPFSPDGKPCTPP ALNCYWPLNDYGFYTTTGIGYQPYRVV VLSFELLNAPATVCGPKLSTDLIKNQCV NFNFNGLTGTGV

SEQ ID NUMBER	Amino Acid Position Relative to SEQ ID NO: 1	Amino Acid Sequence Amino acid residues in bold (amino acid residues 1-16) identify a signal sequence.
53	17-757 plus an N-terminal mouse K chain leader sequence and a C-terminal myc epitope and a poly histidine tag	METDTLLLWVLLLWVPGSTGDDRCTTF DDVQAPNYTQHTSSMRGVYYPDEIFRSD TLYLTQDLFLPFYSNVTGFHTINHTFGNP VIPFKDGIYFAATEKSNVVRGWVFGSTM NNKSQSVIIINNSTNVVIRACNFELCDNPF FAVSKPMGTQTHTMIFDNAFNCTFEYIS DAFSLDVSEKSGNFKHLREFVFKNKDGF LYVYKGYQPIDVVRDLPSGFNTLKPIFKL PLGINITNFRAILTAFSPAQDIWGTSAAA YFVGYLKPTTFMLKYDENGTITDAVDCS QNPLAELKCSVKSFEIDKGIYQTSNFRVV PSGDVVRFPNITNLCPFGEVFNATKFPSV YAWERKKISNCVADYSVLYNSTFFSTFK CYGVSATKLNDLCFSNVYADSFVVKGD DVRQIAPGQTGVIADYNYKLPDDFMGC VLAWNTRNIDATSTGNYNYKYRYLRHG KLRPFERDISNVPFSPDGKPCTPPALNCY WPLNDYGFYTTTGIGYQPYRVVVLSFEL LNAPATVCGPKLSTDLIKNQCVNFNFNG LTGTGVLTPSSKRFQPFQQFGRDVSDFT DSVRDPKTSEILDISPCAFGGVSVITPGTN ASSEVAVLYQDVNCTDVSTAIHADQLTP AWRIYSTGNNVFQTQAGCLIGAEHVDTS YECDIPIGAGICASYHTVSLLRSTSQKSIV AYTMSLGADSSIAYSNNTIAIPTNFSISITT EVMPVSMAKTSVDCNMYICGDSTECAN LLLQYGSFCTQLNRALSGIAAEQEQKLIS EEDLHHHHHH
54	17-276 plus an N-terminal mouse K chain leader sequence and a C-terminal myc epitope and a poly histidine tag	METDTLLLWVLLLWVPGSTGDDRCTTF DDVQAPNYTQHTSSMRGVYYPDEIFRSD TLYLTQDLFLPFYSNVTGFHTINHTFGNP VIPFKDGIYFAATEKSNVVRGWVFGSTM NNKSQSVIIINNSTNVVIRACNFELCDNPF FAVSKPMGTQTHTMIFDNAFNCTFEYIS DAFSLDVSEKSGNFKHLREFVFKNKDGF LYVYKGYQPIDVVRDLPSGFNTLKPIFKL PLGINITNFRAILTAFSPAQDIWGTSAAA YFVGYLKPTTFMLKYDENGTITDAVEQK LISEEDLHHHHHH

SEQ ID NUMBER	Amino Acid Position Relative to SEQ ID NO: 1	Amino Acid Sequence Amino acid residues in bold (amino acid residues 1-16) identify a signal sequence.
55	17-537 plus an N-terminal mouse K chain leader sequence and a C-terminal myc epitope and a poly histidine tag	METDTLLLWVLLLWVPGSTGDDRCTTF DDVQAPNYTQHTSSMRGVYYPDEIFRSD TLYLTQDLFLPFYSNVTGFHTINHTFGNP VIPFKDGIYFAATEKSNVVRGWVFGSTM NNKSQSVIIINNSTNVVIRACNFELCDNPF FAVSKPMGTQTHTMIFDNAFNCTFEYIS DAFSLDVSEKSGNFKHLREFVFKNKDGF LYVYKGYQPIDVVRDLPSGFNTLKPIFKL PLGINITNFRAILTAFSPAQDIWGTSAAA YFVGYLKPTTFMLKYDENGTITDAVDCS QNPLAELKCSVKSFEIDKGIYQTSNFRVV PSGDVVRFPNITNLCPFGEVFNATKFPSV YAWERKKISNCVADYSVLYNSTFFSTFK CYGVSATKLNDLCFSNVYADSFVVKGD DVRQIAPGQTGVIADYNYKLPDDFMGC VLAWNTRNIDATSTGNYNYKYRYLRHG KLRPFERDISNVPFSPDGKPCTPPALNCY WPLNDYGFYTTTGIGYQPYRVVVLSFEL LNAPATVCGPKLSTDLIKNQCVNFNFNG LTGTGV EQKLISEEDLHHHHHH

SEQ ID NUMBER	Amino Acid Position Relative to SEQ ID NO: 1	Amino Acid Sequence Amino acid residues in bold (amino acid residues 1-16) identify a signal sequence.
56	17-756 N-terminal without a signal peptide	DRCTTFDDVQAPNYTQHTSSMRGVYYP DEIFRSDTLYLTQDLFLPFYSNVTGFHTI NHTFGNPVIPFKDGIYFAATEKSNVVRG WVFGSTMNNKSQSVIIINNSTNVVIRACN FELCDNPFFAVSKPMGTQTHTMIFDNAF NCTFEYISDAFSLDVSEKSGNFKHLREFV FKNKDGFLYVYKGYQPIDVVRDLPSGFN TLKPIFKLPLGINITNFRAILTAFSPAQDIW GTSAAAYFVGYLKPTTFMLKYDENGTIT DAVDCSQNPLAELKCSVKSFEIDKGIYQ TSNFRVVPSGDVVRFPNITNLCPFGEVFN ATKFPSVYAWERKKISNCVADYSVLYNS TFFSTFKCYGVSATKLNDLCFSNVYADS FVVKGDDVRQIAPGQTGVIADYNYKLPD DFMGCVLAWNTRNIDATSTGNYNYKYR YLRHGKLRPFERDISNVPFSPDGKPCTPP ALNCYWPLNDYGFYTTTGIGYQPYRVV VLSFELLNAPATVCGPKLSTDLIKNQCV NFNFNGLTGTGVLTPSSKRFQPFQQFGR DVSDFTDSVRDPKTSEILDISPCAFGGVS VITPGTNASSEVAVLYQDVNCTDVSTAI HADQLTPAWRIYSTGNNVFQTQAGCLIG AEHVDTSYECDIPIGAGICASYHTVSLLR STSQKSIVAYTMSLGADSSIAYSNNTIAIP TNFSISITTEVMPVSMAKTSVDCNMYICG DSTECANLLLQYGSFCTQLNRALSGIAA E
57	272-537	ITDAVDCSQNPLAELKCSVKSFEIDKGIY QTSNFRVVPSGDVVRFNITNLCPFGEVFN ATKFPSVYAWERKKISNCVADYSVLYNS TFFSTFKCYGVSATKLNDLCFSNVYADS FVVKGDDVRQIAPGQTGVIADYNYKLPD DFMGCVLAWNTRNIDATSTGNYNYKYR YLRHGKLRPFERDISNVPFSPDGKPCTPP ALNCYWPLNDYGFYTTTGIGYQPYRVV VLSFELLNAPATVCGPKLSTDLIKNQCV NFNFNGLTGTGV
58	24-39 D24 peptide	DVQAPNYTQH TSSMRGC
59	540-555 P540 peptide	PSSKRFQPFQQFGRDC

SEQ ID NUMBER	Amino Acid Position Relative to SEQ ID NO: 1	Amino Acid Sequence Amino acid residues in bold (amino acid residues 1-16) identify a signal sequence.
60	1-16 spike signal sequence	**MFIFLLFLTLTSGSDL**
61	303-537 containing the receptor binding domain	SNFRVVPSGDVVRFPNITNLCPFGEVFNA TKFPSVYAWERKKISNCVADYSVLYNST FFSTFKCYGVSATKLNDLCFSNVYADSF VVKGDDVRQIAPGQTGVIADYNYKLPD DFMGCVLAWNTRNIDATSTGNYNYKYR YLRHGKLRPFERDISNVPFSPDGKPCTPP ALNCYWPLNDYGFYTTTGIGYQPYRVV VLSFELLNAPATVCGPKLSTDLIKNQCV NFNFNGLTGTGV
62	319-517 containing the receptor binding domain	ITNLCPFGEVFNATKFPSVYAWERKKISN CVADYSVLYNSTFFSTFKCYGVSATKLN DLCFSNVYADSFVVKGDDVRQIAPGQT GVIADYNYKLPDDFMGCVLAWNTRNID ATSTGNYNYKYRYLRHGKLRPFERDISN VPFSPDGKPCTPPALNCYWPLNDYGFYT TTGIGYQPYRVVVLSFELLNAPATVCGP KLST
63	319-518 containing the receptor binding domain	ITNLCPFGEVFNATKFPSVYAWERKKISN CVADYSVLYNSTFFSTFKCYGVSATKLN DLCFSNVYADSFVVKGDDVRQIAPGQT GVIADYNYKLPDDFMGCVLAWNTRNID ATSTGNYNYKYRYLRHGKLRPFERDISN VPFSPDGKPCTPPALNCYWPLNDYGFYT TTGIGYQPYRVVVLSFELLNAPATVCGP KLSTD

Example 6

Structure of the Spike Protein

To characterize the properties and function of the SARS-CoV S protein, nucleic acids encoding the full-length Tor2 isolate were cloned into expression vectors as described above. The Tor2 isolate is further described in Marra et al. *The genome sequence of the SARS-associated coronavirus*, Science 300:1399–1404 (2003). Clones generated included the full-length S protein (1255 residues), the ectodomain Se (residues 17–1189) having just the extracellular domain of the S protein with the putative transmembrane domain and cytoplasmic tail of the spike protein deleted, fragments containing the N-

terminal 276 (SEQ ID NO:50), 537 (SEQ ID NO:52), and 756 (SEQ ID NO: 56) amino acid residues (S276, S537, and S756, respectively) including a putative 16-residue signal sequence or a mouse k chain leader sequence, and an internal fragment (S272–537) containing residues 272–537 (SEQ ID NO:57) (see Fig. 1B).

Amino acid residues 758–761 (RNTR) form part of the following general motif for cleavage by precursor convertases:

K/R–Xn–K/R

where X is any amino acid residue and n = 0, 2, 4 or 6.

The S1 subunit is approximately encompassed within the S756 fragment. This finding is in agreement with the size of the S1 subunit for murine coronaviruses, e.g., strain JHM where S1 is 769 residues, and for the human coronavirus OC43 (778 residues). *See* Gallagher & Buchmeier, *Coronavirus spike proteins in viral entry and pathogenesis*, Virology 279: 371–374 (2001); Kunkel & Herrler, *Structural and functional analysis of the surface protein of human coronavirus OC43*, Virology 195 417: 195–202 (1993). However, for the human coronavirus 229E, S1 is considered to consist of a shorter 547 residue fragment that corresponds to S537. Bonavia et al., *Identification of a receptor-binding domain of the spike glycoprotein of human coronavirus HCoV-229E*, J. Virol. 77: 2530–2538 (2003).

All S glycoprotein fragments and the full-length S glycoprotein ran on SDS–PAGE gels at positions significantly higher than their estimated molecular weights, indicating that these polypeptides are likely post-translationally modified. The S276 polypeptide had an apparent molecular weight of about 75 kDa, S537 had an apparent molecular weight of about 100–110 kDa, S756 had an apparent molecular weight of about 130–140 kDa, and Se and S had apparent molecular weights of about 200 kDa or higher (Figs. 4 and 6). The bands corresponding to these polypeptide were broad even when observed at low exposure (Fig. 6; some data not shown). These data indicate significant glycosylation as observed for the S glycoprotein and fragments thereof. Based on approximate estimations of molecular weight it appears that the S2 subunit is not as heavily glycosylated as S756 (constituting the S1 subunit). Notably, S276 is heavily glycosylated if one assumes that only glycosylation contributes to the increased molecular mass.

Most of the SARS-CoV S glycoprotein obtained from cell culture supernatants was not cleaved, although weak bands due to smaller proteins were observed on SDS-PAGE gels. One of these weak bands runs at the same position as S756, suggesting the possibility of inefficient cleavage (Figs. 4 and 6). Random digestion by proteases may occur and further studies are needed to determine if the S glycoprotein cleavage is necessary for its function.

Example 7

Expression of peptide fragments in Escherichia coli

A nucleic acid segment encoding a SEQ ID NO:51 peptide fragment containing amino acid residues 17-446 of SEQ ID NO: 1 was cloned into the pRSET vector (Invitrogen, San Diego, CA) to create the plasmid pRSET-S(17-446). *E. coli* BL21DE3 cells were transformed with pRSET-S(17-446) and then induced with IPTG. The results of the induction are shown in Fig. 2.

Example 8

Use of the T7 promoter to drive expression of a cloned peptide fragment of the invention

Human 293 cells or Monkey Vero E6 cells were grown to a density of $1.2X10^6$ cells/T25 flask (60 mm dish) in 5 ml of DMEM+10% FBS medium the day prior to transfection. The cells were then transfected, using the Polyfect (Qiagen) transfection kit according to the manufacturer's protocol, with pSecTag2B constructs (6 ug each) containing inserts coding for the various peptide fragments of the spike protein. These constructs were prepared as described above.

After 4 hour of transfection, a VTF7.3 vaccinia virus carrying a T7 polymerase was used to infect the transfected cells at a MOI (multiplicity of infection) of 20 (Fuerst et al., Proc. Natl. Acad. Sci., 93:11371 (1986)). This procedure provided for the use of the T7 promoter in the pSecTag2B vector instead of the CMV promoter, which is much weaker (Nussbaum et al., J. Virol., 68:5411 (1994)). After three hours of infection, 1.5 ml of fresh medium was added to the cells and then the cells were transferred to a 31°C incubator. The cells were incubated for an additional 24 hours, after which the culture medium was collected.

No measurable cytopathicity was observed in cells transfected with any of the S nucleic acid constructs (data not shown), indicating that the full-length and soluble fragments of the S glycoprotein may not have significant cytotoxic effects. However, at higher levels of expression such effects are possible and formation of syncytia as described below may lead to cell death.

Example 9

Spike-Specific Antibodies

New Zealand rabbits were immunized with 0.1 mg of various peptides selected by a computer program for their immunogenicity. Serum from the immunized rabbits was tested in ELISA and Western blot for reactivity. Sera from rabbits immunized with two peptides exhibited the highest and specific activity against the spike glycoprotein and were selected for further study. Antibodies denoted D24 and P540 were elicited by the peptides DVQAPNYTQH TSSMRGC (SEQ ID NO:58) and PSSKRFQPFQQFGRDC (SEQ ID NO:59), respectively. Another anti-SARS-CoV S glycoprotein polyclonal antibody IMG-542, which recognizes amino acid 288–303 of the S glycoprotein, was purchased from IMGENEX (San Diego, CA).

Example 10

Immunoprecipitation and Purification of Spike Polypeptides

Soluble spike polypeptides fragments were obtained from the Vero E6 or 293 cell culture medium. However, the full-length spike glycoprotein was detected only in the cell lysate.

Medium from cells transfected with nucleic acids encoding various soluble S fragments was collected and subjected to centrifugation at 1000g for 10 min to remove cellular debris. The cleared medium was incubated with either Ni–NTA agarose beads (Qiagen, Valencia, CA) or an immunoprecipitating antibody plus glycoprotein G–Sepharose beads (Sigma, St. Louis, MO) for 2 h at 4 °C. The beads were then mixed with an equal volume of SDS gel sample buffer, boiled for 3 min, and subjected to gel analysis. For full-length S glycoprotein, cells were lysed first in PBS supplemented with 1% NP-40 and 0.5mM PMSF for 1 h at 4 °C, and centrifuged at 14,000 rpm in a table-top

Eppendorf centrifuge for 20 min. The cleared lysate was either immunoprecipitated first or used directly in Western blotting.

Example 11

Western blotting and slot blots

Cells expressing the S glycoprotein were lysed first with a PBS-based NP40 lysis buffer as described above, and the debris was cleared by centrifugation. For soluble S fragments the medium was collected and cleared as described above. For slot blots, the cleared lysate or medium from supernatant was used directly to blot the nitrocellulose membrane following the protocol suggested by the manufacturer (Bio-Rad, Hercules, CA) and the membrane was subjected to antibody detection as in conventional Western blotting. For Western blotting, a monoclonal anti-c-Myc epitope antibody (Invitrogen, Carlsbad, CA) or anti-spike protein rabbit polyclonal antibodies obtained by immunization of rabbits with spike peptides were diluted in TBST buffer. Antibodies were incubated with the membrane for 2 h, washed and then the membrane was incubated with a secondary antibody conjugated with HRP for 1 h, washed four times (each time for 15 min), and then developed using the ECL reagent (Pierce, Rockford, IL).

Example 12

Cell-binding assay and ELISA

Medium containing soluble S fragments was collected and cleared by centrifugation. Vero E6 or other cells ($5x10^6$) were incubated with 0.5 ml of cleared medium containing soluble S fragments and 2 μg of anti-c-Myc epitope antibody conjugated with HRP at 4 °C for 2 h. Cells were then washed three times with ice-cold PBS and collected by centrifugation. The cell pellets were incubated with ABTS substrate from Roche (Indianapolis, IN) at RT for 10 min, the substrate was cleared by centrifugation, and the optical density at 405 nm was measured. The result of the slot blot analysis is presented in Fig. 4 and discussed in further detail below.

For ELISA, purified ACE2 (R&D, Minneapolis, MN) was adsorbed onto Maxisorp ELISA plates in pH 9.6 buffer at a concentration of 100 ng per well. Medium 154 (150 μl) containing various soluble S fragments and 0.6 μg of anti-

c-155 Myc epitope antibodies conjugated with HRP were incubated in each well at 37 °C for 2 h. Wells were washed and 60 μl of ABTS substrate was added to each well. The optical density (OD_{405}) was measured 20 min later.

Example 13

Fluorescent dye redistribution cell fusion assay

HeLa or 293T cells, transfected with plasmids encoding the S glycoprotein, were loaded with Calcein AM (Molecular Probes), which is converted within the cells to calcein green. The cells were incubated in medium containing 1 μg/ml Calcein AM for 1 h at 37 °C and 5% CO2, and then washed and re-suspended in fresh medium. Plated target cells, Vero E6, were stained with CMAC (Molecular Probes) by incubation in 1 μg/ml CMAC in medium for 30 min at 37 °C and 5% CO_2. The cells were then washed twice with medium, incubated for 20 min in fresh medium, washed again, and then covered with 0.5 ml medium per well. The S-expressing cells, loaded with calcein, were added to the target cells and incubated for 1, 2, or 4 h at 37 °C and 5% CO_2. Fusion was measured as the ratio between the cells that have double staining and the total number of target cells in contact with an S glycoprotein-expressing cell. Microphotographs were taken using the MethaMorph 4.0 software from Universal Imaging.

Example 14

β-Galactosidase reporter gene-based cell–cell fusion assay

293T cells ($1.5x10^6$) were plated in T25 flasks. The next day, these cells were separately transfected with pCDNA3-S, pSectag2B-S, pCDNA3-ACE2, and pCDNA3-ACE2-Ecto using the Polyfect transfection kit (Qiagen, Valencia, CA) following the manufacturer's suggested protocol. Four hours after transfection, cells transfected with S constructs were infected with T7 polymerase-expressing vaccinia virus VTF7.3 and cells transfected with ACE-2 constructs were infected with β-gal encoding vaccinia virus (VCB21R). Two hours after infection, cells were incubated with fresh medium and transferred to 31 °C for overnight incubation. The next day S glycoprotein-expressing cells and ACE-2-expressing cells were mixed in a 1:1 ratio and incubated at 37 °C. Three hours later, cells were lysed by adding NP-40 to a final concentration of 0.5%.

Cell lysate (50 μl) was mixed with equal volume of CPRG substrate and OD_{595} was measured 1 hr later.

Example 15

Expression of Spike Polypeptides in Mammalian Cells

For certain experiments, all proteins except the full-length S glycoprotein were tagged with a c-Myc epitope and a histidine tag. These proteins were expressed in 293 and Vero E6 cells after transfection with the corresponding plasmids followed by infection with vaccinia virus-expressing T7 polymerase.

The tagged proteins were detected by using an anti-c-Myc monoclonal antibody (Fig. 4). As shown in Fig. 4, the T7 promoter was a highly efficient promoter for expression of the S glycoprotein. In these experiments, the T7 promoter gave rise to higher levels of expression than the CMV promoter, which under most circumstances is a strong promoter (Fig. 4A). As shown in Fig. 4A, the S fragments were soluble and their concentration in the culture supernatants was inversely proportional to their size.

Example 16

Anti-Spike Antibodies

To be able to detect unlabeled proteins, validate the data obtained by the anti-c-Myc antibody, and localize possible antigenic sites rabbit polyclonal antibodies were developed. Two of these antibodies, D24 and P540, were raised against peptides starting at residues 24 and 540, respectively. The D24 and P540 antibody preparations specifically recognized certain soluble fragments (Fig. 4C). As expected, D24 recognized all fragments; P540 recognized S756, Se, and S but not the smaller fragments (Fig. 4C; some data not shown). The D24 antibody preparation was relatively weak. However, the P540 preparation was very sensitive even at dilution 1:10,000 and was used extensively in the experiments described herein.

The P540 antibody preparation was used to detect whether the S glycoprotein was expressed intracellularly, extracellularly or on the cell surface. As shown in Fig. 5, the full-length S glycoprotein was expressed at the cell surface, although at low levels, as measured by flow cytometry.

Example 17

Spike Protein Mediates Cell Fusion

The full-length S glycoprotein mediates fusion at neutral pH with cells expressing receptor molecules. Cell–cell fusion assays were performed to confirm that the full-length recombinant S glycoprotein was functional, and to ascertain whether the S protein requires other viral proteins and/or low pH for its fusion activity.

Expression of the full-length S glycoprotein with both vectors pCDNA3-S and pSectag2B-S, supported fusion with ACE2 expressing cells efficiently, as evidenced by formation of syncytia of various sizes and by β-gal reporter gene-based assay (Fig. 7). Interestingly, the pSectag2B-S construct in which the S glycoprotein leader peptide was replaced by a mouse k chain leader sequence induced faster formation of syncytia. Moreover, the syncytia formed were larger and more numerous than those induced by pCDNA3-S, which encodes the native S glycoprotein (data not shown). The extent of fusion mediated by S expressed from pSectag2B-S was also higher than from pCDNA3-S as measured by a reporter gene-based assay (Fig. 7B). These data indicate that the natural S glycoprotein may not be efficiently transported to the cell surface. These studies also suggest that the β-gal assay described here can serve as a quick and quantitative method to identify inhibitors of SAR-CoV entry into cells, as well as a tool to study SARS-CoV entry mechanism.

Notably, fusion of Vero E6 cells was not detected using the β-gal assay or the syncytium formation assay when the cells were not transfected with plasmids encoding ACE2 and the cells expressed only native concentrations of the receptor. To explore the possibility that this was due to low sensitivity of these two assays, another assay was used. This new assay was based on fluorescent dye redistribution that is able to detect fusion of single cells. Even with this fluorescent-based assay statistically significant differences between cells transfected with plasmids encoding the full-length S glycoprotein and various negative controls were not detected. Some of the negative controls included transfection with plasmids encoding soluble S fragments at different pH (data not shown). Significant cell-cell fusion was only detected when the cells were transfected with plasmids encoding ACE2, suggesting that the higher levels of receptor expression achieved by expression of recombinant ACE2 could be

important for cell–cell fusion. Overall, these results suggest that recombinant S glycoprotein can mediate cell fusion, that fusion can occur at neutral pH, and that its efficiency is dependent on the concentration of the receptor molecules.

Moreover, soluble fragments of the S glycoprotein inhibit S-mediated cell fusion. As shown in Fig. 15, addition of S fragments S272-537 and S17-537, which have the receptor binding domain as described below, inhibit S-mediated cell fusion. In this assay, the S272-537 (SEQ ID NO:57) fragment, exhibited the most inhibition. The S17-276 fragment that does not have the receptor binding domain exhibited little or no inhibition of S-mediated cell fusion. These data indicate that S polypeptide fragments that have the receptor binding domain could inhibit SARS-CoV fusion with animal cells, thereby inhibiting or preventing SARS-CoV infection.

Hence, blocking, modulating or inhibiting the activity of the spike protein receptor binding domain, with an anti-RBD antibody, S polypeptide, S peptide or aptamer may be an effective preventive or treatment for SARS-CoV infection.

Example 18

Identification of Spike Protein Receptor-Binding Domain

This Example illustrates that the Spike protein receptor-binding domain is localized within residues 272 to 537 (SEQ ID NO:57), and likely within residues 303-537 (SEQ ID NO:61). Later experiments have shown that a fragment containing residues 319-517 (SEQ ID NO:62) also has receptor binding activity.

An assay based on the binding of various soluble fragments to receptor expressing Vero E6 cells was developed to localize the receptor-binding domain (RBD) of the S glycoprotein. This assay involved measurement of fluorescence associated with binding of antibodies directed against the S polypeptides to Vero E6 cells and was developed prior to the identification of the SARS-CoV receptor. Vero E6 cells that are susceptible to SARS-CoV infection were incubated with full-length S polypeptide and various soluble S fragments. Several cell lines that are not susceptide to SARS-CoV infection were similarly incubated with full-length S polypeptides and soluble fragments thereof.

As shown in Figs. 8A and 8B, all fragments S fragments bound to Vero E6 cells except the smallest one S fragment (S276). No such binding was detected when several cell lines that are not susceptide to SARS-CoV infection were incubated with full-length S polypeptides and soluble fragments thereof. Binding to Vero E6 cells was proportional to the expression levels of the fragments and was approximately inversely proportional to the sizes of the fragments. These findings suggested that the RBD is localized between residues 272 and 537.

To further localize the RBD, an antibody (IMG 542) was used that was generated using a peptide containing residues 288–303. This antibody did not inhibit binding of the S537 fragment to Vero E6 cells although it did bind to the S537 fragment (Fig. 8B; some data not shown), suggesting that the RBD is localized between residues 303 and 537. Because of the relatively large antibody size and the possibility for steric hindrance, it is likely that the RBD is downstream of residue 303. Recently, the RBD of the HCoV-229E was localized to a fragment containing amino acid residues 407–547. Ksiazek et al. *A novel coronavirus associated with severe acute respiratory syndrome,* N. Engl. J. Med. 348: 1953–1966 (2003); Rota et al. *Characterization of a novel coronavirus associated with severe acute respiratory syndrome,* Science 300: 1394–1399 (2003). In contrast, the RBD for murine hepatitis virus was mapped to the N-terminal 330 amino acids.

It remains to be seen whether there is structural similarity between the RBD-containing fragments of the SARS-CoV S1 glycoprotein (e.g., S272–537) and the HCoV-229E or hepatitis virus RBD, and whether such similarity is related to the use of the same host for replication. These two viruses use different receptors. The straightforward cell-binding approach described here could also be helpful for identification of other virus receptors.

Recently, workers have reported the identification of ACE2 as a functional receptor for the SARS-CoV. Li et al. *Angiotensin-converting enzyme 2 is a functional receptor for the SARS coronavirus,* Nature 426: 450-54 (2003). The identification of ACE2 as receptor permitted further validation that the results provided above are correct. As shown in Fig. 8C, when purified ACE2 is used in an ELISA to test for binding, the same binding pattern was observed as

for the cell-binding assay. This was true for all of the S fragments tested (Fig. 8C).

The results provided herein not only offer new tools to study entry of the SARS virus into cells, confirm that ACE2 is a receptor for the SARS-CoV S1 glycoprotein and localize the RBD but also facilitate development of novel vaccine immunogens and therapeutics for prevention and treatment of SARS.

Example 19

N-terminal and C-terminal Oligomerization of the S glycoprotein

This Example illustrates that the extreme N-terminal fragment of the S glycoprotein, upstream from the RBD, may play a role in fusion, and the S ectodomain forms trimers that could mediate fusion through six-helix bundle intermediates.

Materials and methods

Antibodies and plasmids. The rabbit anti-S serum used in Western and FACS analyses, P540 was developed by the inventors as described above. See also, Xiao et al. Biochem. Biophys. Res. Comm. 312: 1159-65 (2003). The anti-Myc epitope antibody was purchased from Invitrogen (Carlsbad, CA). The anti-ACE2 goat polyclonal antibody was purchased from R&D system (Minneapolis, MN) and used for detection by Western blotting.

Site directed mutagenesis was used to create the consensus cleavage sites corresponding to that of the HIV-1 envelope glycoprotein (Env) and some coronaviruses within the full length SARS-CoV S glycoprotein gene in pCDNA3. The QuickChange Kit from Stratagene (La Jolla, CA) was employed using the protocol provided by manufacturer. For expression of various N terminal S fragments, the corresponding gene fragments were amplified by PCR and cloned into the pSecTag2 expression vector (Invitrogen, Carlsbad). The plasmid pCDNA3-ACE2-ecto, which expresses the ACE2 soluble ectodomain tagged with C9 peptide was kindly provided by Michael Farzan (Harvard University, Boston MA).

Protein expression and purification. Various N terminal fragments of the S glycoprotein were sub-cloned in pSecTag2 expression vector and transfected into 293T cells followed by infection with VTF7.3 as described in Xiao et al.

Biochem. Biophys. Res. Comm. 312: 1159-65 (2003). The protein expressed and secreted into the medium was purified using the HiTrap Ni^{++}-Chelating column (Pharmacia) under native conditions. The purified protein was dialyzed against PBS buffer and stored for further analysis.

S glycoprotein dimerization and its interaction with ACE2 examined by co-immunoprecipitation. For S fragment dimerization, different S glycoprotein constructs, alone or in combination, were transfected to 293T cells as described in Xiao et al. Biochem. Biophys. Res. Comm. 312: 1159-65 (2003). Medium containing S fragments was subjected to immunoprecipitation with rabbit anti-S polyclonal antiserum P540. For some co-immunoprecipitation experiments, DTT was added to create reducing condition to eliminate inter-molecule interactions through disulfide bonds. Immunoprecipitated S fragments were detected by Western using an anti-Myc epitope monoclonal antibodies. Soluble ACE2-C9 was expressed similarly. ACE2-C9 secreted into the medium was used directly for incubation with various S fragments for 2 hours at 4°C. Afterwards, ACE2 was immunoprecipitated by incubating with 1D4 anti-C9 monoclonal antibody and protein G-Sepharose beads at 4°C for one hour. The beads were washed four times with PBS, suspended in SDS-PAGE sample buffer, boiled for 3 min and subjected to gel separation. The presence of either ACE2 or S in the sample was examined by Western as described in Xiao et al. Biochem. Biophys. Res. Comm. 312: 1159-65 (2003).

Flow cytometry. Cells transfected with full length S glycoprotein or S glycoprotein with different N terminal deletions and infected with VTF7.3 were incubated with the P540 rabbit anti-S polyclonal antibody and goat anti-rabbit antibody conjugated with FITC in PBS containing 1% BSA at 4°C for two hours. Cells were then washed four times in ice cold PBS and analyzed with FacsCalibur (Becton Dickinson, San Jose, California).

Gel filtration analysis of S fragments. After being purified on Ni-chelating column and buffer-exchanged to PBS, S fragment samples were loaded onto a Superose 12 10/300 GL column (Pharmacia, Uppsala, Sweden) that had been pre-equilibrated with PBS. The proteins were eluted with PBS at 0.5 ml/min, and 0.5 ml fractions were collected. The Superose 12 column was calibrated with protein molecular mass standard of 669, 440, 232, 158, 67, 44

and 25 kD. A 10 μl aliquot was taken from each fraction for Western blot analysis.

Crosslinking. Purified S537 fragment was diluted to a concentration of 0.2 μg/ml in PBS. BS^3 (Pierce, Rockford, IL) was added to the S537 solution to a final concentration of 1 mg/ml and incubated on ice for 1 min. The samples were then mixed with an equal volume of 4X SDS-PAGE loading buffer and analyzed by Western blot.

Cell fusion β-gal reporter gene assay. Cells transfected with pSecTag2B-S or pCDNA3-ACE2 and infected with VTF7.3 and VCB21R respectively were collected by trypsin digestion and washed once with PBS. Cells were then suspended in regular DMEM medium at pH 7.4 and mixed. Cells were lysed after four hours of incubation and β-gal activity was measured using CPRG as the substrate (Roche) as described in Xiao et al. Biochem. Biophys. Res. Comm. 312: 1159-65 (2003).

ELISA. Two ELISA assays were used. In the sandwich ELISA the plate was coated with an anti-His tag antibody, then the S fragment were added and detected with an anti-c-Myc epitope antibody. This assay was used for detection of the S fragments. In the second ELISA assay the C9 tagged receptor ACE2 was coated on the plates through an anti-C9 antibody (ID4) and the S fragments were added and after washing detected with an anti-c-Myc epitope antibody. In all experiments the incubations with the c-Myc epitope antibody were for 2 hours at room temperature. The optical density (OD) was measured and normalized to the highest value.

Results

The N-terminal fragment upstream of the RBD of the S glycoprotein forms a dimer. It has been previously shown for another coronavirus (MHV) that soluble S1 (similar to SU) fragments form dimers, that the extreme N-terminal 330 amino acid residue region that contains the receptor binding domain participates in the dimerization, and that only dimers bind the receptor CEACAM. *See* Lewicki & Gallagher, J. Biol. Chem. 277:19727-34 (2002). However, the inventors and others have localized the SARS-CoV receptor binding domain downstream from the extreme N-terminus. Xiao et al. Biochem.

Biophys. Res. Comm. 312: 1159-65 (2003); Wong et al. J. Biol. Chem. 279: 3197-3201 (2004); Babcock et al. J. Virol. 78: 4552-4560 (2004).

To address the possibility of oligomerization by receptor binding domain-containing fragments and to assess their function in mediating membrane fusion, several S fragments were tested for oligomerization. These S fragments included the extreme N-terminal fragment (residues 17 through 276 denoted as S276, SEQ ID NO:50) that does not bind the receptor ACE2, several S fragments (S756, S537, S272-537) that bind ACE2, as well as a fragment including residues 319 through 517 (denoted as S319-517, SEQ ID NO:62) that retains receptor binding activity. These fragments were selected in part because they fold independently and are secreted in the cell culture supernatant, although the efficiency of their expression varied significantly (Fig. 9A, left) and their concentration was decreased when co-expressed with S756 (Fig. 9A, right).

To find whether any of these fragments oligomerizes with the largest one (S756) that includes the equivalent of the receptor-binding subunit of the envelope glycoproteins (SU in general and S1 for coronaviruses) the polypeptide fragments were coexpressed, and then the mixtures in the cell culture supernatants were immunoprecipitated with the antibody P540. As described in previous Examples, this rabbit polyclonal antibody preparation was developed against a peptide containing residues 540-555 (SEQ ID NO:59) of the S glycoprotein. The P450 antibody binds the S756 polypeptide but not the other fragments (Fig. 9B, left). All N-terminal fragments except the smallest fragment (S319-517) containing the receptor binding domain were coimmunoprecipated with S756 by P540 (Fig. 9B, right). To rule out the possibility of nonspecific disulfide bond formation that may lead to coimmunoprecipitation, DTT was included in one of the coimmunoprecipitation experiments. DTT had no effect on either immunoprecipitation or coimmunoprecipitation of secreted S756 (left lanes) or S756+S276 (right lanes) (Fig. 9C, left panel).

To find the size of the oligomers, one of the fragments (S537) was cross-linked with BS^3. The right panel of Fig. 9C shows the appearance of a new band with a molecular weight corresponding to a dimer but not of higher order oligomers. To exclude the possibility of artifacts due to cross-linking and further to confirm the formation of dimers, the S537 fragment was also analyzed by gel filtration. Two gel filtration elution peaks were observed: one due to species of

molecular weight of about 230 kDa and the other one of about 110 kDa (Fig. 10A, upper panel) corresponding to a dimer-sized oligomer and a monomer, respectively. In contrast, the smallest fragment containing the receptor binding domain (S319-517) was eluted only as a monomer at about 35 kDa molecular weight (Fig. 2A, lower panel). Overall, these results suggest that soluble SU is a dimer and that the dimerization domain is within the extreme N-terminal region upstream from residue 317 and the receptor binding domain.

The dimeric N terminal region is required for S mediated cell-cell fusion. Because the putative dimerization domain is upstream from the receptor binding domain within S1 and the fusion machinery is in S2, one might hypothesize that dimerization may not be required for mediation of fusion. To test this hypothesis, two deletion mutants of the full-length S glycoprotein were generated. The N-terminal 103 residues were deleted from one fragments and the N-terminal 311 residues were deleted from another (Fig. 9A), thereby eliminating the presumed dimerization domain. Both mutants did not exhibit any fusion activity compared to the wild type full-length S glycoprotein, which did (Fig. 9A). To test whether a differential level of expression could account for the lack of observable fusogenic activity, the surface and overall levels of expression were measured by flow cytometry and Western blotting. The data from both assays suggested that the level of expression of the two deletion mutants is undistinguishable from that of the wild type (Figs. 11B and C). These results suggest that the extreme N-terminus is required for fusion by a mechanism that may or may not involve dimerization.

Dimeric S1 binds ACE2 much more efficiently than monomeric fragments containing the Receptor Binging Domain. Previous work with another coronavirus (MHV) suggested that only dimeric S1 binds its receptor CEACAM. Lewicki & Gallagher, J. Biol. Chem. 277:19727-734 (2002). Experiments were conducted on SARS-CoV fragments to understand how the dimeric state of the S1 may affect fusion. In particular, binding of S1 fragments in monovalent and bivalent form to ACE2 was observed by using the anti-c-Myc epitope antibody for conversion of monovalent S1 fragments into bivalent ones. One of these S1 fragments (S319-517, SEQ ID NO:62) did not bind to any measurable degree to surface-immobilized ACE2 unless bound by an anti-c-Myc epitope antibody, which converted it into a bivalent molecule in solution before

and during incubation with the receptor (Fig. 12). In contrast, S537 bound to ACE2 without the antibody although the antibody presence increased its binding (Fig. 12). These results suggest that a dimeric state of S1 could contribute to an increased overall affinity that may enhance fusion efficiency.

The soluble S ectodomain is a trimer. Viral envelope glycoproteins of class I fusion proteins such as hemagglutinin (HA) of influenza are trimeric through the transmembrane domain. Because the SARS-CoV S glycoprotein was recently found to be class I fusion protein, the S2 subunit may facilitate trimerization of the whole S glycoprotein. However, a dimeric S1 with a trimeric S2 could lead to higher order oligomers whose formation depends on the availability of the dimerization binding site in the native S glycoprotein. To test this possibility the size of the soluble S ectodomains (Se) was approximated by gel filtration, where the transmembrane domain and the cytoplasmic tail were deleted. As shown in Fig. 13, a complex having the approximate size of a trimer (MW 512 kDa) was detected. No higher order oligomers were detected. These results not only suggest that the Se fragment and perhaps the full-length membrane-associated S are trimers in there native unbound state but also indicate that the dimerization site in S1 is not readily available for intertrimer interactions.

These results indicate the following: 1) the SU subunit of the SARS-CoV S glycoprotein (S1) forms dimers, 2) the dimerization domain does not overlap and is upstream of the receptor binding domain, 3) deletion of the dimerization domain abolishes fusion, 4) dimeric S1 binds receptor molecules much more efficiently than monovalent fragments containing the receptor binding domain, and 5) the soluble S ectodomain forms trimers under gel filtration conditions.

It has been previously reported that some SU subunits of class I fusion proteins (that bind receptor molecules) can form dimers including, for example, gp120 of the retrovirus HIV-1 and S1 of the coronavirus MHV. Center et al. J. Virol. 74: 4448-55 (2000); Lewicki et al. J. Biol. Chem. 277: 19727-34 (2002). Until the present work, the role of S1 dimerization for mediation of membrane fusion was unclear. It is now generally accepted that soluble ectodomains such as the gp140 protein of the HIV-1 and SIV envelope glycoproteins (Env) form trimers although dimers and tetramers can be observed. Center et al. Proc. Nat'l

Acad. Sci. U.S.A. 98: 14877-82 (2001). Similarly, it appears that at least a possible fusion intermediate quaternary structure of coronaviruses including the SARS-CoV of S2 is trimeric. Liu et al. Lancet 363: 938-947 (2004); Bosch et al. Proc. Nat'l Acad. Sci. U.S.A. 101: 8455-60 (2004). In contrast, some data indicates that the MHV S2 protein is monomeric after dissociation from S1. Lewicki et al. J. Biol. Chem. 277: 19727-34 (2002). Dimer-to-trimer transitions play a critical role in the mechanism of fusion mediated by class II fusion proteins. Thus it has been proposed that changes in the quaternary structure of some coronaviruses may play a role in the fusion mechanism. *Id.* One should note that both the HIV-1 Env and the MHV S glycoproteins are cleaved and the SU can dissociate from the transmembrane subunit, however, such dissociation may not be important for fusion. In contrast, the SARS-CoV S is not cleaved when expressed in membrane associated or soluble form and cleavage may not be required for fusion. Thus, although the SARS-CoV S glycoprotein is a class I fusion protein, the lack of cleavage is an exception from the rule that the Envs of class I fusion proteins are cleaved presumably to confer a metastable high-energy state that could drive the fusion reaction.

This finding that the SU (S1) domain of the SARS-CoV S glycoprotein can form dimers and also forms trimers with the ectodomain of the transmembrane domain (S2) poses an interesting topological situation. Thus, if two of the monomers within a trimer also form a dimer, then the third monomer would still be free to interact with a "free" monomer from another trimer and form a dimer of the two trimers. In another scenario the orientation of each of the monomers in the trimer may not allow formation of dimers in the trimer but leave "free" binding sites for dimerization with monomers from other trimers. In this case one might expect the formation of a network of trimers. Finally, the three-dimensional structure of the trimer may not allow any interactions of the monomer dimerization sites with other monomers in the same or different trimer. The later possibility is supported by the preliminary data provided herein where higher order oligomers were not detected using the described gel filtration conditions. Under those conditions either intratrimer dimerization occurs but the third monomer conformation does not allow interactions with monomers from other trimers or such interactions are too weak to be detected, or the trimer three-dimensional structure is such that it does not allow dimerization interactions.

Data provided herein demonstrate lack of fusion after deletion of portions of the dimerization domain and indicate that the dimerization region may play a role in fusion although its mechanism may not be through dimerization interactions. In addition, under native conditions where the surface concentration of the S glycoprotein can be very high, as seen in electron micrographs, it is possible that dimerization interactions play a role in stabilizing a "network" of interacting molecules perhaps somewhat similar to networks of proteins that mediate entry of class II fusion proteins. Such networks, if any, could increase the avidity of interaction with receptor molecules and perhaps facilitate the formation of the fusion pore structure by providing a pre-assembled network of Env molecules or even provide energy to drive the fusion reaction in the absence of S cleavage that generates a high-energy metastable state.

Example 20

Sera from Mice Immunized with DNA Encoding RBD Polypeptides Inhibits S-Mediated Cell Fusion

This Example illustrates that immunizing mammals with DNA encoding receptor binding domain polypeptides may prevent SARS infection.

Materials and Methods

Mice were divided into three groups: group A of mice # 1 through 5 were immunized with plasmid pSecTag-SRBD that encodes for the S319-518 fragment that includes the receptor binding domain (RBD) of the spike protein; group B of mice #1 to #5 were immunized with the plasmid pEAK-10-RBD-Fc that encodes for a fusion protein of RBD (S319-518) fragment fused to Fc and group C mice #1 to #3 which were immunized with a control plasmid. Five BALB/C mice per group were immunized at day 0, day 14 and day 28. Mice received less than 2 ug DNA per immunization with a gene gun. Sera were collected at day 56. In Fig. 14A-B, the first number denotes an individual mouse, the letter denotes the respective immunization group, and the last number denotes the dilution used.

Cells (293T) were incubated with anti-sera from the immunized mice and then mixed with cells expressing S protein. Fusion was measured as described in previous Examples (see also, Xiao et al. BBRC 2003). PC denotes positive

control where no serum was added. For mice #1 to #2 in each group, serum dilution factors of 10, 100, and 1000 were used. For mice #3-#5 in groups A and B, and #3 in the control group, dilution factors of 20 and 100 were used.

Results

The antibody titers for the anti-sera obtained from the mice are shown in Fig. 14A. As shown, mice immunized with DNA encoding the spike protein receptor binding domain (S319-518, groups A and B) had very high titer anti-sera – dilutions up to 1:7250 still reacted strongly to antigen in ELISA assays.

As shown in FIG. 14B, anti-sera from mice immunized with DNA encoding the spike protein receptor binding domain inhibited fusion of cells that express the S protein in a dose dependent manner. Thus, anti-sera from mouse 1A and 2A, which were immunized with DNA encoding the S receptor binding domain, substantially eliminated S-protein mediated cell fusion when used at a 1:10 dilution. Higher dilutions (1:100 and 1:1000) of this anti-sera were less effective. Similar results were observed on cell fusion inhibited by anti-sera from mouse 3A (1:20 dilution), from mouse 4A (1:20 dilution), and from mouse 5A (1:20 dilution).

These data indicate that immunizing mammals with DNA encoding S protein receptor binding domain polypeptides can raise a strong immune response against the spike protein and could prevent SARS infection. As described above, soluble fragments of the S glycoprotein that have the receptor binding domain inhibit S-mediated cell fusion (see Fig. 15).

References

D.S. Dimitrov, Cell biology of virus entry, Cell 101 (2000) 697–702.

D.S. Dimitrov, Virus entry: molecular mechanisms and biomedical applications, Nat. Rev. Microbiol. 2 (2004) 109-122.

K.V. Holmes, SARS-associated coronavirus, N. Engl. J. Med. 348 (2003) 1948–1951.

M.M. Lai, D. Cavanagh, The molecular biology of coronaviruses, Adv. Virus Res. 48 (1997) 1–100.

T.G. Ksiazek, D. Erdman, C.S. Goldsmith, S.R. Zaki, T. Peret, S. Emery, S. Tong, C. Urbani, J.A. Comer, W. Lim, P.E. Rollin, S.F. Dowell, A.E.

Ling, C.D. Humphrey, W.J. Shieh, J. Guarner, C.D. Paddock, P. Rota, B. Fields, J. DeRisi, J.Y. Yang, N. Cox, J.M. Hughes, J.W. LeDuc, W.J. Bellini, L.J. Anderson, A novel coronavirus associated with severe acute respiratory syndrome, N. Engl. J. Med. 348 (2003) 1953–1966.

P.A. Rota, M.S. Oberste, S.S. Monroe, W.A. Nix, R. Campagnoli, J.P. Icenogle, S. Penaranda, B. Bankamp, K. Maher, M.H. Chen, S. Tong, A. Tamin, L. Lowe, M. Frace, J.L. DeRisi, Q. Chen, D.Wang, D.D. Erdman, T.C. Peret, C. Burns, T.G. Ksiazek, P.E. Rollin, A. Sanchez, S. Li.ck, B. Holloway, J. Limor, K. McCaustland, M. Olsen-Rasmussen, R. Fouchier, S. Gunther, A.D. Osterhaus, C. Drosten, M.A. Pallansch, L.J. Anderson, W.J. Bellini, Characterization of a novel coronavirus associated with severe acute respiratory syndrome, Science 300 (2003) 1394–1399.

M.A. Marra, S.J. Jones, C.R. Astell, R.A. Holt, A. Brooks-Wilson, Y.S. Butterfield, J. Khattra, J.K. Asano, S.A. Barber, S.Y. Chan, A. Cloutier, S.M. Coughlin, D. Freeman, N. Girn, O.L. Gri.th, S.R. Leach, M. Mayo, H. McDonald, S.B. Montgomery, P.K. Pandoh, A.S. Petrescu, A.G. Robertson, J.E. Schein, A. Siddiqui, D.E. Smailus, J.M. Stott, G.S. Yang, F. Plummer, A. Andonov, H. Artsob, N. Bastien, K. Bernard, T.F. Booth, D. Bowness, M. Czub, M. Drebot, L. Fernando, R. Flick, M. Garbutt, M. Gray, A. Grolla, S. Jones, H. Feldmann, A. Meyers, A. Kabani, Y. Li, S. Normand, U. Stroher, G.A. Tipples, S. Tyler, R. Vogrig, D. Ward, B. Watson, R.C. Brunham, M. Krajden, M. Petric, D.M. Skowronski, C. Upton, R.L. Roper, The genome sequence of the SARS-associated coronavirus, Science 300 (2003) 1399–1404.

T.M. Gallagher, M.J. Buchmeier, Coronavirus spike proteins in viral entry and pathogenesis, Virology 279 (2001) 371–374.

F. Kunkel, G. Herrler, Structural and functional analysis of the surface protein of human coronavirus OC43, Virology 195 (1993) 195–202.

A. Bonavia, B.D. Zelus, D.E. Wentworth, P.J. Talbot, K.V. Holmes, Identi.cation of a receptor-binding domain of the spike glycoprotein of human coronavirus HCoV-229E, J. Virol. 77 (2003) 2530–2538.

W. Li, M.J. Moore, N. Vasilieva, J. Sui, S.K. Wong, M.A. Berne, M. Somasundaran, J.L. Sullivan, K. Luzuriaga, T.C. Greenough, H. Choe,

M. Farzan, Angiotensin-converting enzyme 2 is a functional receptor for the SARS coronavirus, Nature 426: 450-54 (2003).

P.L. Earl, B. Moss, Mutational analysis of the assembly domain of the HIV-1 envelope glycoprotein, AIDS Res.Hum.Retroviruses. 9 (1993) 589-94.

R.J. Center, P. Schuck, R.D. Leapman, L.O. Arthur, P.L. Earl, B. Moss, J. Lebowitz, Oligomeric structure of virion-associated and soluble forms of the simian immunodeficiency virus envelope protein in the prefusion activated conformation, Proc.Natl.Acad.Sci.U.S.A 98 (2001) 14877-14882.

R.J. Center, P.L. Earl, J. Lebowitz, P. Schuck, B. Moss, The human immunodeficiency virus type 1 gp120 V2 domain mediates gp41-independent intersubunit contacts, J.Virol. 74 (2000) 4448-4455.

D.N. Lewicki, T.M. Gallagher, Quaternary structure of coronavirus spikes in complex with carcinoembryonic antigen-related cell adhesion molecule cellular receptors, J. Biol. Chem. 277 (2002) 19727-19734.

B. Tripet, M.W. Howard, M. Jobling, R.K. Holmes, K.V. Holmes, R.S. Hodges, Structural characterization of the SARS-coronavirus spike S fusion protein core, J Biol. Chem. 279 (2004) 20836-20849.

J. Zhu, G. Xiao, Y. Xu, F. Yuan, C. Zheng, Y. Liu, H. Yan, D.K. Cole, J.I. Bell, Z. Rao, P. Tien, G.F. Gao, Following the rule: formation of the 6-helix bundle of the fusion core from severe acute respiratory syndrome coronavirus spike protein and identification of potent peptide inhibitors, Biochem.Biophys.Res.Commun. 319 (2004) 283-288.

B.J. Bosch, B.E. Martina, Z.R. van der, J. Lepault, B.J. Haijema, C. Versluis, A.J. Heck, R. de Groot, A.D. Osterhaus, P.J. Rottier, Severe acute respiratory syndrome coronavirus (SARS-CoV) infection inhibition using spike protein heptad repeat-derived peptides, Proc. Natl. Acad. Sci. U.S.A 101 (2004) 8455-8460.

S. Liu, G. Xiao, Y. Chen, Y. He, J. Niu, C.R. Escalante, H. Xiong, J. Farmar, A.K. Debnath, P. Tien, S. Jiang, Interaction between heptad repeat 1 and 2 regions in spike protein of SARS-associated coronavirus: implications for virus fusogenic mechanism and identification of fusion inhibitors, Lancet 363 (2004) 938-947.

B.J. Bosch, Z.R. van der, C.A. de Haan, P.J. Rottier, The coronavirus spike protein is a class I virus fusion protein: structural and functional characterization of the fusion core complex, J.Virol. 77 (2003) 8801-8811.

H. Bisht, A. Roberts, L. Vogel, A. Bukreyev, P.L. Collins, B.R. Murphy, K. Subbarao, B. Moss, Severe acute respiratory syndrome coronavirus spike protein expressed by attenuated vaccinia virus protectively immunizes mice, Proc. Natl. Acad. Sci. U.S.A 101 (2004) 6641-6646.

G.J. Babcock, D.J. Esshaki, W.D. Thomas, Jr., D.M. Ambrosino, Amino acids 270 to 510 of the severe acute respiratory syndrome coronavirus spike protein are required for interaction with receptor, J.Virol. 78 (2004) 4552-4560.

G. Simmons, J.D. Reeves, A.J. Rennekamp, S.M. Amberg, A.J. Piefer, P. Bates, Characterization of severe acute respiratory syndrome-associated coronavirus (SARS-CoV) spike glycoprotein-mediated viral entry, Proc.Natl.Acad.Sci.U.S.A 101 (2004) 4240-4245.

S.K. Wong, W. Li, M.J. Moore, H. Choe, M. Farzan, A 193-amino acid fragment of the SARS coronavirus S protein efficiently binds angiotensin-converting enzyme 2, J.Biol.Chem. 279 (2004) 3197-3201.

All patents and publications referenced or mentioned herein are indicative of the levels of skill of those skilled in the art to which the invention pertains, and each such referenced patent or publication is hereby incorporated by reference to the same extent as if it had been incorporated by reference in its entirety individually or set forth herein in its entirety. Applicants reserve the right to physically incorporate into this specification any and all materials and information from any such cited patents or publications.

The specific methods and compositions described herein are representative of preferred embodiments and are exemplary and not intended as limitations on the scope of the invention. Other objects, aspects, and embodiments will occur to those skilled in the art upon consideration of this specification, and are encompassed within the spirit of the invention as defined by the scope of the claims. It will be readily apparent to one skilled in the art

that varying substitutions and modifications may be made to the invention disclosed herein without departing from the scope and spirit of the invention. The invention illustratively described herein suitably may be practiced in the absence of any element or elements, or limitation or limitations, which is not specifically disclosed herein as essential. The methods and processes illustratively described herein suitably may be practiced in differing orders of steps, and that they are not necessarily restricted to the orders of steps indicated herein or in the claims. As used herein and in the appended claims, the singular forms "a," "an," and "the" include plural reference unless the context clearly dictates otherwise. Thus, for example, a reference to "a host cell" includes a plurality (for example, a culture or population) of such host cells, and so forth. Under no circumstances may the patent be interpreted to be limited to the specific examples or embodiments or methods specifically disclosed herein. Under no circumstances may the patent be interpreted to be limited by any statement made by any Examiner or any other official or employee of the Patent and Trademark Office unless such statement is specifically and without qualification or reservation expressly adopted in a responsive writing by Applicants.

The terms and expressions that have been employed are used as terms of description and not of limitation, and there is no intent in the use of such terms and expressions to exclude any equivalent of the features shown and described or portions thereof, but it is recognized that various modifications are possible within the scope of the invention as claimed. Thus, it will be understood that although the present invention has been specifically disclosed by preferred embodiments and optional features, modification and variation of the concepts herein disclosed may be resorted to by those skilled in the art, and that such modifications and variations are considered to be within the scope of this invention as defined by the appended claims.

The invention has been described broadly and generically herein. Each of the narrower species and subgeneric groupings falling within the generic disclosure also form part of the invention. This includes the generic description of the invention with a proviso or negative limitation removing any subject

matter from the genus, regardless of whether or not the excised material is specifically recited herein.

Other embodiments are within the following claims. In addition, where features or aspects of the invention are described in terms of Markush groups, those skilled in the art will recognize that the invention is also thereby described in terms of any individual member or subgroup of members of the Markush group.

WHAT IS CLAIMED:

1. A polypeptide fragment of SEQ ID NO: 1, or a conservative variant thereof, wherein the polypeptide can produce a humoral or cellular immune response when used to inoculate an animal.

2. A polypeptide having any one of SEQ ID NOs: 13, 14, 15, 20-59, 61-63, wherein the polypeptide, wherein the polypeptide can produce a humoral or cellular immune response when used to inoculate an animal.

3. A polypeptide having any one of SEQ ID NOs: 13, 14, 15, 25, 34, 46, 51, 52, 56, 57, 58, 59, 61, 62 or 63, wherein the polypeptide, wherein the polypeptide can produce a humoral or cellular immune response when used to inoculate an animal.

4. The polypeptide of claim 1, 2 or 3, wherein the polypeptide is soluble in an aqueous solution.

5. The polypeptide of claim 1, 2 or 3, wherein the animal is a mammal.

6. The polypeptide of claim 5, wherein the mammal is a human.

7. The polypeptide of claim 1, 2 or 3, wherein the polypeptide is amino-terminally or carboxyl-terminally blocked.

8. A coupled protein comprising a carrier protein coupled to a second polypeptide having any one of (a) SEQ ID NOs: 13, 14, 15, 20-59, 61-63; (b) a peptide fragment of SEQ ID NO: 1, or a conservative variant of (a) or (b).

9. The coupled protein of claim 8, wherein the carrier protein is soluble in an aqueous solution.

10. The coupled protein of claim 9, wherein the carrier protein is selected from the group consisting of bovine serum albumin, keyhole limpet hemacyanin, ovalbumin, mouse serum albumin, rabbit serum albumin.

11. The coupled protein of claim 8, wherein the coupled protein produces a humoral or a cellular immune response when used to inoculate an animal.

12. The coupled protein of claim 11, wherein the animal is a mammal.

13. The coupled protein of claim 12, wherein the mammal is a human.

14. An immunopeptide comprising a polypeptide having any one of (a) SEQ ID NOs: 13, 14, 15, 20-59, 61-63; or (b) a fragment of SEQ ID NO: 1; coupled to arsanilic acid, sulfanilic acid, an acetyl group, or a picryl group.

15. The immunopeptide of claim 14, wherein the immunopeptide produces a humoral or a cellular immune response when used to inoculate an animal.

16. The immunopeptide of claim 15, wherein the animal is a mammal.

17. The immunopeptide of claim 16, wherein the mammal is a human.

18. An immune composition comprising an adjuvant and a polypeptide having any one of SEQ ID NOs: 13, 14, 15, 20-59, 61-63; or a fragment of SEQ ID NO: 1.

19. The immune composition of claim 18, wherein the adjuvant is selected from the group consisting of aluminum hydroxide, lipid A, killed bacteria, polysaccharide, mineral oil, Freund's incomplete adjuvant, Freund's complete adjuvant, aluminum phosphate, iron, zinc, a calcium salt, acylated tyrosine, an acylated sugar, a cationically derivatized polysaccharide, an anionically derivatized polysaccharide, a

polyphosphazine, a biodegradable microsphere, a monophosphoryl lipid A, and quil A.

20. The immune composition of claim 18, wherein the polypeptide is amino-terminally or carboxyl-terminally blocked.

21. A peptidomimetic of an amino acid sequence having any one of (a) SEQ ID NOs: 13, 14, 15, 20-59, 61-63; (b) a fragment of SEQ ID NO: 1, or a conservative variant of (a) or (b).

22. An immune composition comprising an adjuvant and a peptidomimetic of an amino acid sequence having any one of (a) SEQ ID NOs: 13, 14, 15, 20-59, 61-63; (b) a fragment of SEQ ID NO: 1; or a conservative variant of (a) or (b).

23. The immune composition of claim 22, wherein the adjuvant is selected from the group consisting of aluminum hydroxide, lipid A, killed bacteria, polysaccharide, mineral oil, Freund's incomplete adjuvant, Freund's complete adjuvant, aluminum phosphate, iron, zinc, a calcium salt, acylated tyrosine, an acylated sugar, a cationically derivatized polysaccharide, an anionically derivatized polysaccharide, a polyphosphazine, a biodegradable microsphere, a monophosphoryl lipid A, and quil A.

24. A nucleic acid segment that encodes a polypeptide having any one of (a) SEQ ID NOs: 13, 14, 15, 20-59, 61-63; (b) a peptide that is a fragment of SEQ ID NO: 1, or a conservative variant of (a) or (b).

25. An expression cassette comprising a promoter that is operably linked to a nucleic acid segment that encodes a polypeptide having any one of (a) SEQ ID NOs: 13, 14, 15, 20-59, 61-63; (b) a peptide that is a fragment of SEQ ID NO: 1; or a conservative variant of (a) or (b).

26. The expression cassette according to claim 25, wherein the promoter is a constitutive promoter or a regulated promoter.

27. A nucleic acid construct comprising a vector and a nucleic acid segment that encodes (a) a polypeptide having any one of SEQ ID NOs: 13, 14, 15, 20-59, 61-63; (b) a peptide that is a fragment of SEQ ID NO: 1; (c) a conservative variant of (a) or (b); or an expression cassette according to claim 25.

28. The nucleic acid construct according to claim 27, wherein the vector is selected from the group consisting of a plasmid, a cosmid, a yeast artificial chromosome, a bacterial artificial chromosome, an F-factor, a virus, an expression vector, and a phagemid.

29. A recombinant virus comprising a viral vector and a nucleic acid segment that encodes (a) a polypeptide having any one of SEQ ID NOs: 13, 14, 15, 20-59, 61-63; (b) a peptide that is a fragment of SEQ ID NO: 1; (c) a conservative variant of (a) or (b); or an expression cassette according to claim 25.

30. The recombinant virus of claim 29, wherein the viral vector is selected from the group consisting of vaccinia virus, canarypox, adenovirus, and herpes virus.

31. A composition comprising a pharmaceutical carrier and (a) a polypeptide having any one of SEQ ID NOs: 13, 14, 15, 20-59, 61-63; (b) a peptide that is a fragment of SEQ ID NO: 1; or (c) a conservative variant of (a) or (b).

32. The composition of claim 31, wherein the composition is formulated for treatment of SARS-CoV.

33. The composition of claim 31, wherein the composition is formulated for inhibition of SARS-CoV fusion with, or entry into, mammalian cells.

34. A composition comprising a pharmaceutical carrier and a nucleic acid segment that encodes (a) a polypeptide having any one of SEQ ID NOs: 13, 14, 15, 20-59, 61-63; (b) a peptide that is a fragment of SEQ ID NO: 1; (c) a conservative variant of (a) or (b); or an expression cassette according to claim 30.

35. The composition of claim 34, wherein the composition is formulated for treatment of SARS-CoV.

36. The composition of claim 34, wherein the composition is formulated for prevention of SARS-CoV fusion with, or entry into, mammalian cells.

37. A viral vaccine comprising a pharmaceutical carrier, a viral vector and a nucleic acid segment that encodes (a) a polypeptide having any one of SEQ ID NOs: 13, 14, 15, 20-59, 61-63; (b) a peptide that is a fragment of SEQ ID NO: 1; (c) a conservative variant of (a) or (b); or an expression cassette according to claim 30.

38. The viral vaccine according to claim 34, wherein the viral vaccine is formulated in unit dosage form.

39. A peptide vaccine comprising a pharmaceutical carrier and (a) a peptide having any one of SEQ ID NOs: 13, 14, 15, 20-59, 61-63; (b) a fragment of SEQ ID NO: 1; (c) a peptidomimetic of (a) or (b); (d) or a conservative variant of (a) or (b).

40. The peptide vaccine according to claim 39, wherein the peptide vaccine is formulated in unit dosage form.

41. A microorganism vaccine comprising a pharmaceutical carrier and a microorganism that expresses (a) a peptide having any one of SEQ ID NOs: 13, 14, 15, 20-59, 61-63; (b) a fragment of SEQ ID NO: 1; or (c) a conservative variant of (a) or (b).

42. The microorganism vaccine according to claim 41, wherein the microorganism is selected from the group consisting of Salmonella and Listeria monocytogenes.

43. The microorganism vaccine according to claim 42, wherein the microorganism vaccine is formulated in unit dosage form.

44. A DNA vaccine comprising a pharmaceutical carrier and vector into which is inserted a nucleic acid segment that encodes (a) an amino acid sequence as put forth in any one of SEQ ID NOs: 13, 14, 15, 20-59, 61-63; (b) a fragment of SEQ ID NO: 1; or (c) a conservative variant of (a) or (b).

45. The DNA vaccine according to claim 44, wherein the vector is selected from the group consisting of a plasmid, a cosmid, a yeast artificial chromosome, a bacterial artificial chromosome, an F-factor, a virus, and a phagemid.

46. The DNA vaccine according to claim 44, wherein the DNA vaccine is formulated in unit dosage form.

47. The DNA vaccine according to claim 46, wherein the DNA vaccine further comprises a myonecrotic agent.

48. The DNA vaccine according to claim 47, wherein the myonecrotic agent is bupivicaine or cardiotoxin.

49. An antibody that binds to an amino acid sequence as set forth in any one of SEQ ID NOs: 13, 14, 15, 20-59, 61-63; or a fragment of SEQ ID NO: 1.

50. The antibody according to claim 49, wherein the antibody specifically binds to an amino acid sequence as set forth in any one of SEQ ID NOs: 13, 14, 15, 20-59, 61-63; or a fragment of SEQ ID NO: 1.

51. The antibody according to claim 49, wherein the antibody specifically binds to a S protein receptor binding domain.

52. The antibody according to claim 49, wherein the antibody is a monoclonal antibody, a polyclonal antibody, a single-chain antibody, an antigen-binding antibody fragment, or a humanized antibody.

53. The antibody according to claim 52, wherein the antigen-binding antibody fragment is an scFv, Fv, Fab', Fab, diabody, linear antibody or $F(ab')_2$.

54. The antibody according to claim 49, wherein the antibody is coupled to a detectable tag.

55. The antibody according to claim 54, wherein the detectable tag is a fluorescent protein, a fluorescent marker, a radiolabel, an enzyme, or an affinity tag.

56. The antibody according to claim 49, wherein the antibody is coupled to a toxin.

57. The antibody according to claim 56, wherein the toxin is an A chain toxin, a ribosome inactivating protein, α-sarcin, gelonin, aspergillin, restrictocin, a ribonuclease, an epipodophyllotoxin, diphtheria toxin, Pseudomonas exotoxin, ricin, doxorubicin, daunorubicin, taxol, ethiduim bromide, mitomycin, etoposide, tenoposide, vincristine, vinblastine, colchicine, dihydroxy anthracin dione, actinomycin D, PE40, abrin, or a glucocorticoid.

58. A pharmaceutical composition comprising a pharmaceutical carrier and an antibody that binds to an amino acid sequence as set forth in any one of SEQ ID NOs: 13, 14, 15, 20-59, 61-63; or a fragment of SEQ ID NO: 1.

59. A method to immunize a mammal against severe acute respiratory syndrome comprising administering to the mammal a therapeutically effective amount of an antibody that binds to an amino acid sequence as set forth in any one of SEQ ID NOs: 13, 14, 15, 20-59, 61-63; or a fragment of SEQ ID NO: 1.

60. The method of claim 59, wherein the antibody specifically binds to an amino acid sequence as set forth in any one of SEQ ID NOs: 13, 14, 15, 20-59, 61-63; or a fragment of SEQ ID NO: 1.

61. The method of claim 59, wherein the mammal is a human.

62. A method to treat severe acute respiratory syndrome in a mammal comprising administering to the mammal a therapeutically effective amount of an antibody that binds to an amino acid sequence as set forth in any one of SEQ ID NOs: 13, 14, 15, 20-59, 61-63; or a fragment of SEQ ID NO: 1.

63. The method of claim 62, wherein the antibody specifically binds to an amino acid sequence as set forth in any one of SEQ ID NOs: 13, 14, 15, 20-59, 61-63; or a fragment of SEQ ID NO: 1.

64. The method of claim 62, wherein the mammal is a human.

65. The method of claim 59 or 62, wherein the antibody is formulated with a pharmaceutical carrier or diluent.

66. A method for treating or inhibiting severe acute respiratory syndrome in a mammal comprising administering to the mammal a therapeutically

effective amount of a S polypeptide comprising an amino acid sequence as set forth in any one of SEQ ID NOs: 13, 14, 15, 20-59, 61-63; or a fragment of SEQ ID NO: 1.

67. A method for raising an immune response in a mammal against a SARS coronavirus spike protein comprising administering a therapeutically effective amount of a polypeptide comprising an amino acid sequence as set forth in any one of SEQ ID NOs: 13, 14, 15, 20-59, 61-63; or a fragment of SEQ ID NO: 1.

68. The method of claim 67, wherein the polypeptide comprises an amino acid sequence as set forth in any one of SEQ ID NOs: 13, 14, 15, 25, 34, 51, 52, 56, 57, 58, 59, 61, 62, 63; or a fragment of SEQ ID NO: 1.

69. The method of claim 67, wherein the mammal is a human.

70. A method to diagnose severe acute respiratory syndrome in an animal comprising:

(a) contacting a biological sample obtained from the animal with an antibody that binds to an amino acid sequence as set forth in any one of SEQ ID NOs: 13, 14, 15, 20-59, 61-63; or a fragment of SEQ ID NO: 1; and

(b) determining if the antibody binds to the biological sample.

71. The method of claim 70, wherein the animal is a mammal.

72. The method of claim 70, wherein the mammal is a human.

73. A method for making an antibody comprising: obtaining an animal that was immunized with (a) a peptide fragment of a polypeptide having an amino acid sequence as set forth in SEQ ID NO: 1; (b) a polypeptide having an amino acid sequence as set forth in any one of SEQ ID NOs: 13, 14, 15, 20-59, 61-63; (c) a peptidemimetic of (a) or (b), or (d) a

conservative variant of (a) or (b); and isolating an antibody that binds to (a).

74. A method to make an antibody comprising: obtaining an animal that was immunized with a coupled protein having a carrier protein coupled to (a) a peptide fragment of a polypeptide having an amino acid sequence as set forth in SEQ ID NO: 1; (b) a polypeptide having an amino acid sequence as set forth in any one of SEQ ID NOs: 1, 13, 14, 15, 20-55; (c) a peptidemimetic of (a) or (b), or (d) a conservative variant of (a) or (b); and isolating an antibody that binds to a polypeptide having an amino acid sequence as set forth in SEQ ID NO:1.

75. A kit comprising packaging material and an antibody or aptamer that binds to an amino acid sequence as set forth in any one of SEQ ID NOs: 1, 13, 14, 15, 20-59, 61-63; or a fragment of SEQ ID NO: 1.

76. The kit of claim 75, wherein the antibody is formulated with a pharmaceutical carrier or diluent.

77. The kit of claim 75, further comprising a syringe.

78. A kit comprising packaging material and a therapeutically effective amount of a S polypeptide comprising an amino acid sequence as set forth in any one of SEQ ID NOs: 13, 14, 15, 20-59, 61-63; or a fragment of SEQ ID NO: 1.

79. The kit of claim 78, wherein the S polypeptide is formulated with a pharmaceutical carrier or diluent.

80. The kit of claim 78, further comprising a syringe.

81. A monoclonal antibody that specifically binds to an amino acid sequence as set forth in any one of SEQ ID NOs:1, 13, 14, 15, 20-59, 61-63.

82. An isolated polyclonal antibody that specifically binds to an amino acid sequence as set forth in any one of SEQ ID NOs:1, 13, 14, 15, 20-59, 61-63.

83. An aptamer that binds to an amino acid sequence as set forth in any one of SEQ ID NOs:1, 13, 14, 15, 20-59, 61-63; or a fragment of SEQ ID NO: 1.

84. A pharmaceutical composition comprising a pharmaceutical carrier and an aptamer that binds to an amino acid sequence as set forth in any one of SEQ ID NOs:1, 13, 14, 15, 20-59, 61-63; or a fragment of SEQ ID NO: 1.


1
2
3

Fig. 1A

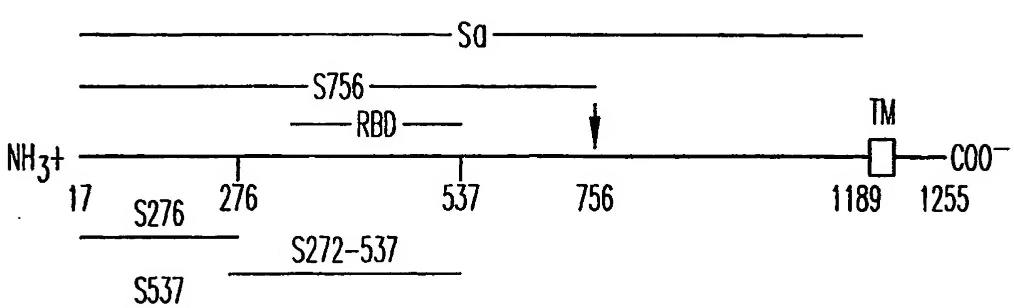

Sa
S756
RBD
TM
NH3+
COO−
17
276
537
756
1189
1255
S276
S272-537
S537

Fig. 1B

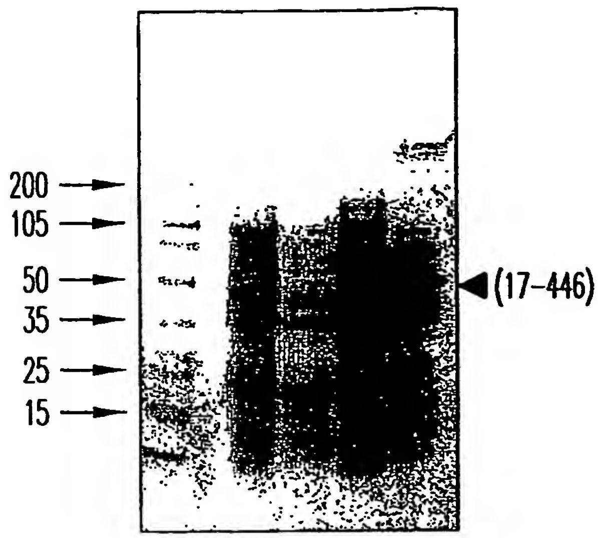

200
105
50
35
25
15
(17-446)

Fig. 2

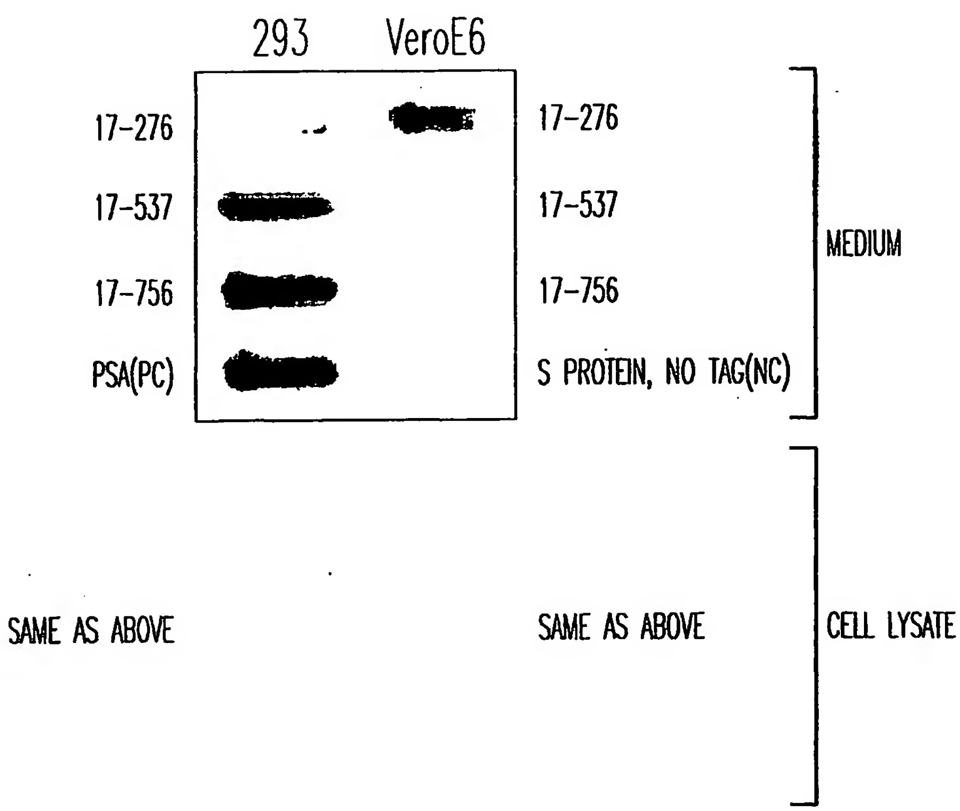

293
VeroE6
17-276
17-537
17-756
PSA(PC)
17-276
17-537
17-756
S PROTEIN, NO TAG(NC)
MEDIUM
SAME AS ABOVE
SAME AS ABOVE
CELL LYSATE

Fig. 3

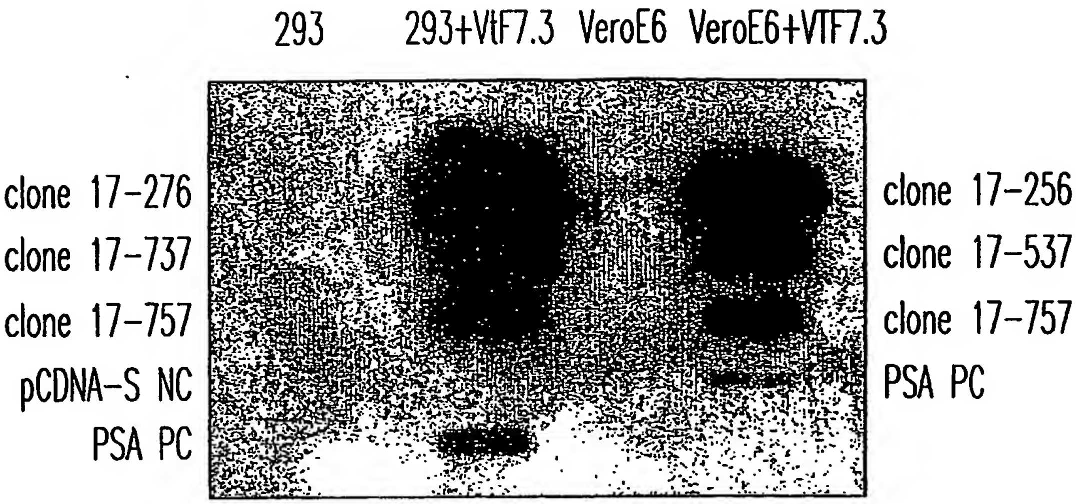

293
293+VtF7.3
VeroE6
VeroE6+VTF7.3
clone 17-276
clone 17-737
clone 17-757
pCDNA-S NC
PSA PC
clone 17-256
clone 17-537
clone 17-757
PSA PC

Fig. 4A

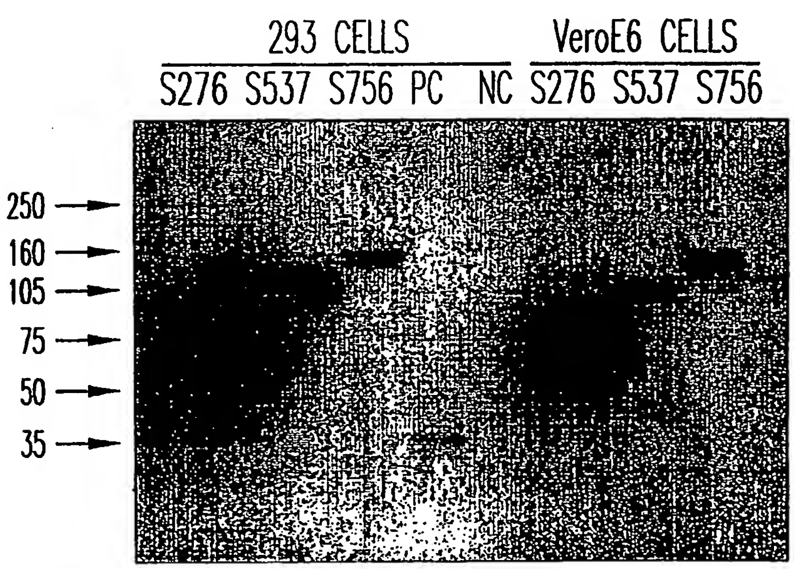

293 CELLS
VeroE6 CELLS
S276
S537
S756
PC
NC
S276
S537
S756
250
160
105
75
50
35

Fig. 4B

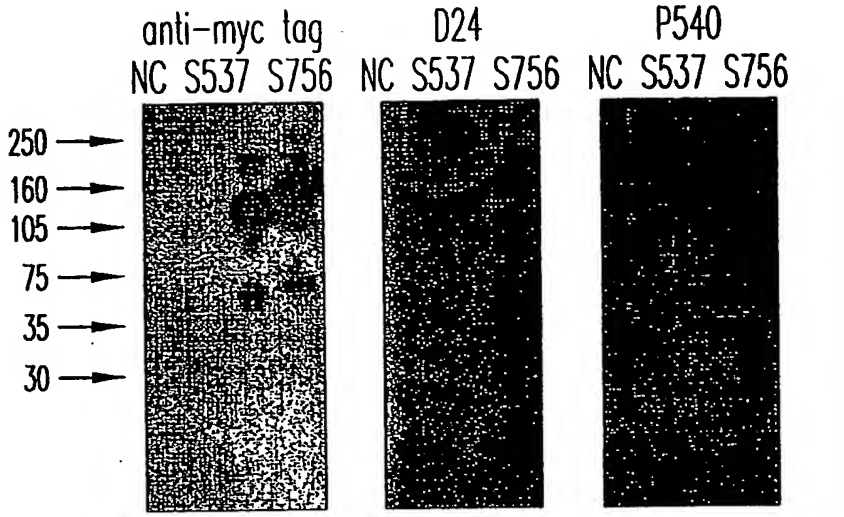

anti-myc tag
NC S537 S756
D24
NC S537 S756
P540
NC S537 S756
250
160
105
75
35
30

Fig. 4C

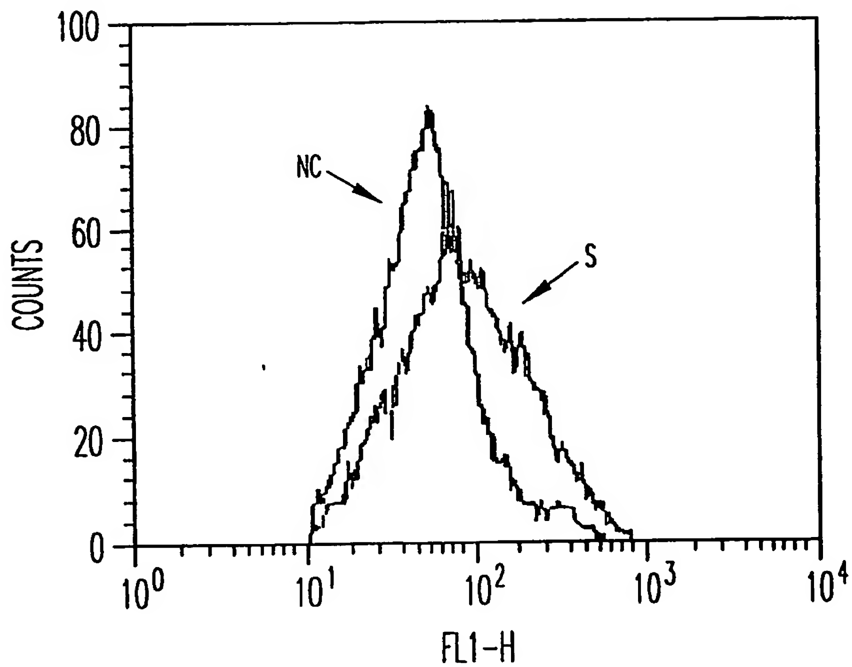

100
80
60
40
20
0
COUNTS
NC
S
10^0
10^1
10^2
10^3
10^4
FL1-H

Fig. 5

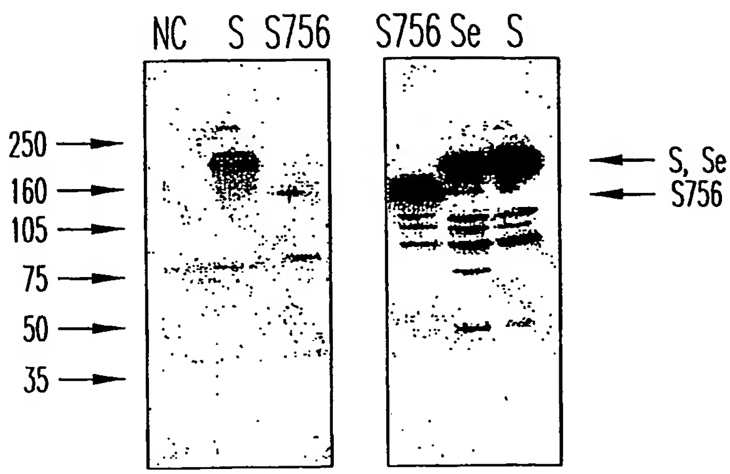


Fig. 6A Fig. 6B

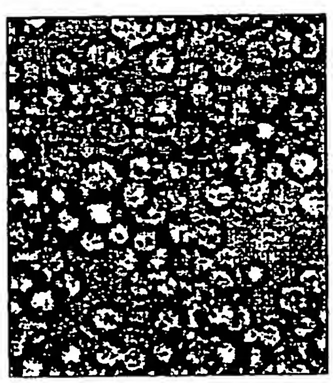


Fig. 7A

Fig. 7B

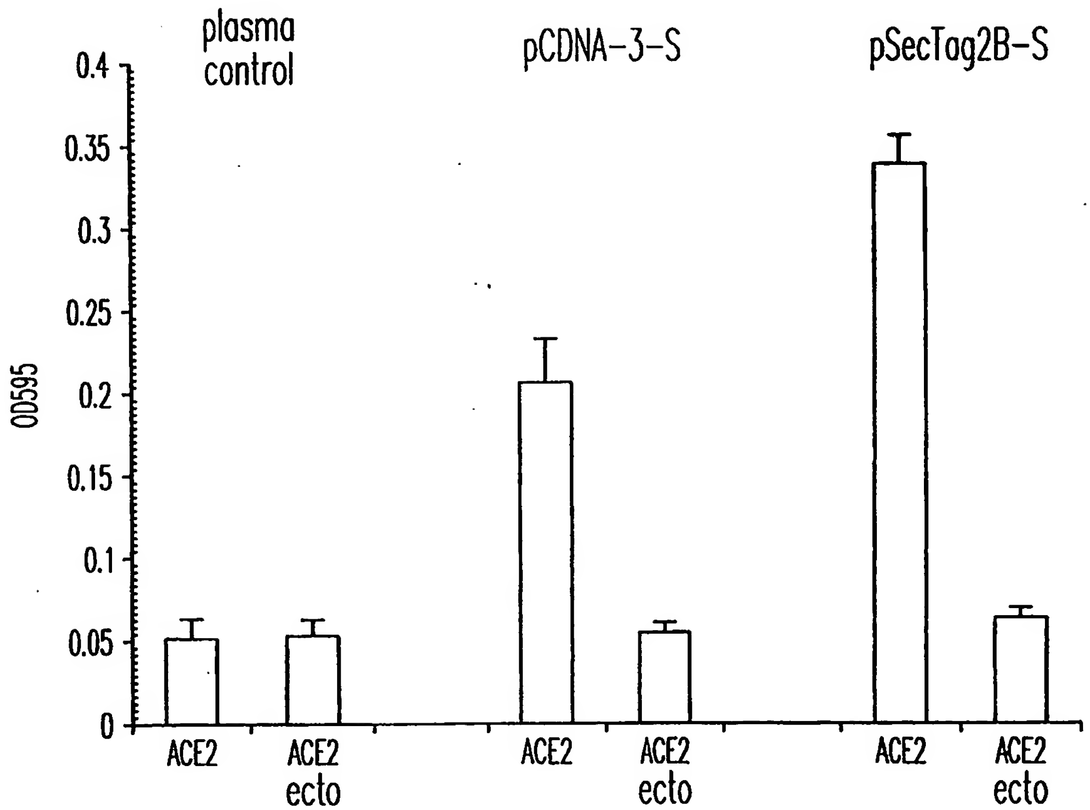

plasma
control
pCDNA-3-S
pSecTag2B-S
0.4
0.35
0.3
0.25
0.2
0.15
0.1
0.05
0
OD595
ACE2
ACE2
ecto
ACE2
ACE2
ecto
ACE2
ACE2
ecto

Fig. 7C

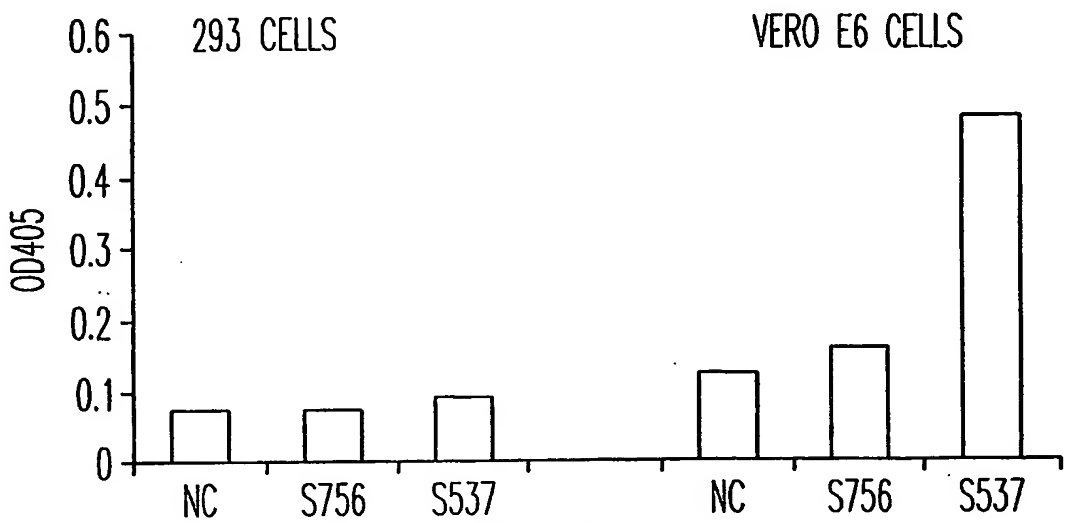


Fig. 8A

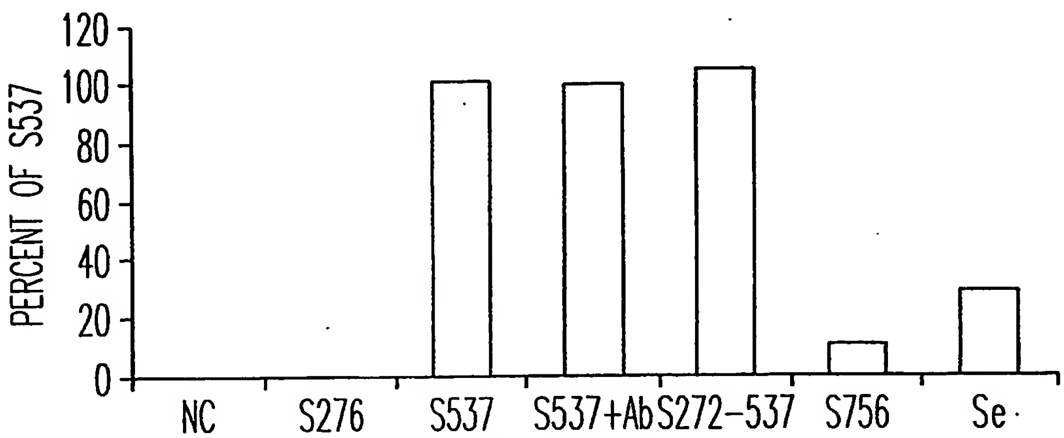


Fig. 8B

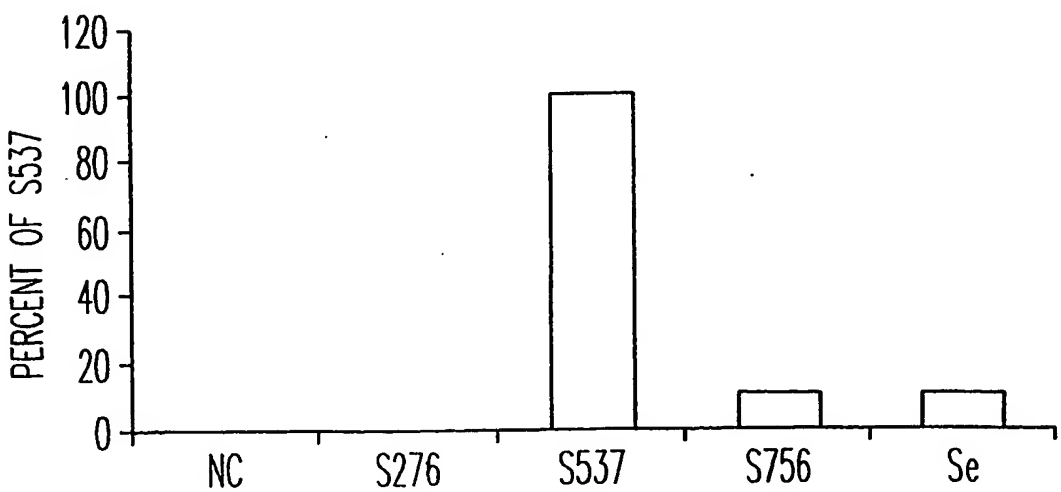


Fig. 8C

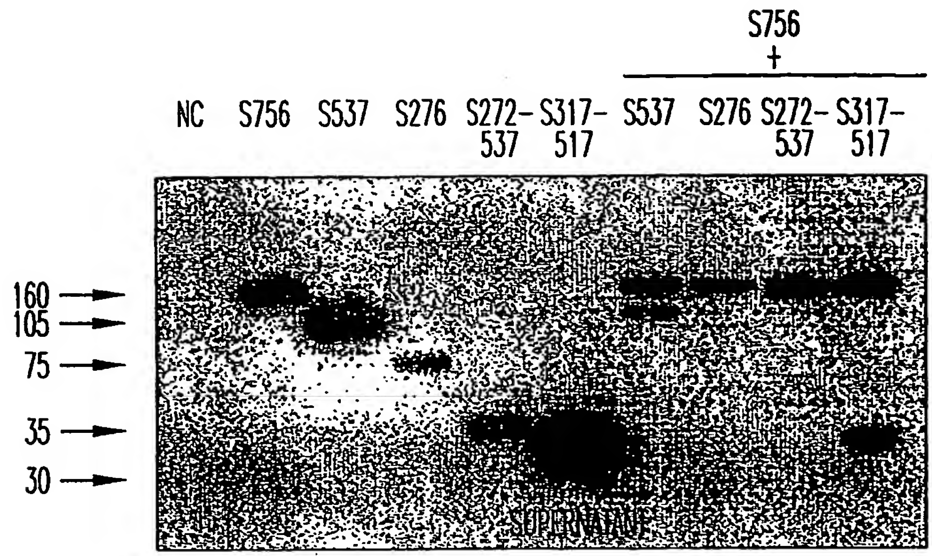


Fig. 9A

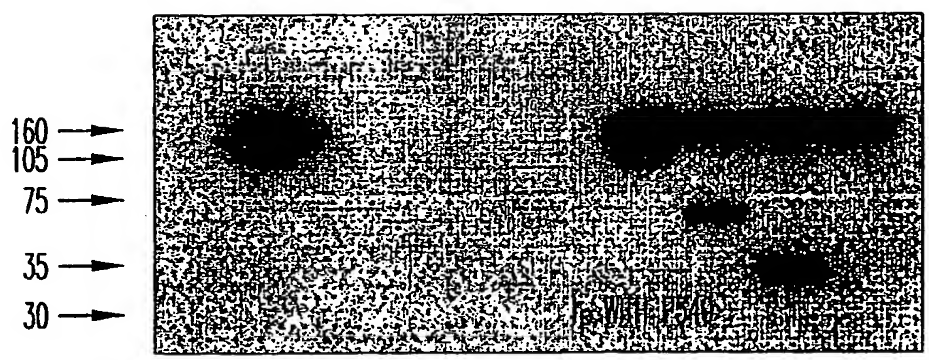


Fig. 9B

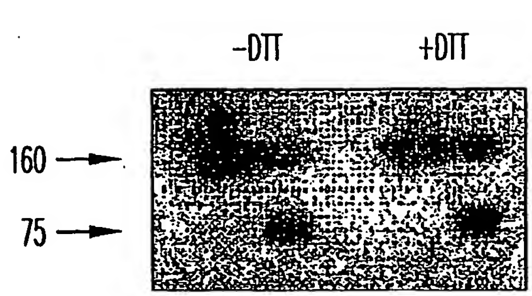


Fig. 9C

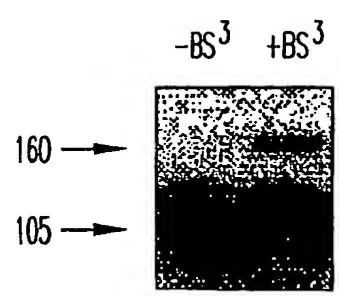


Fig. 9D

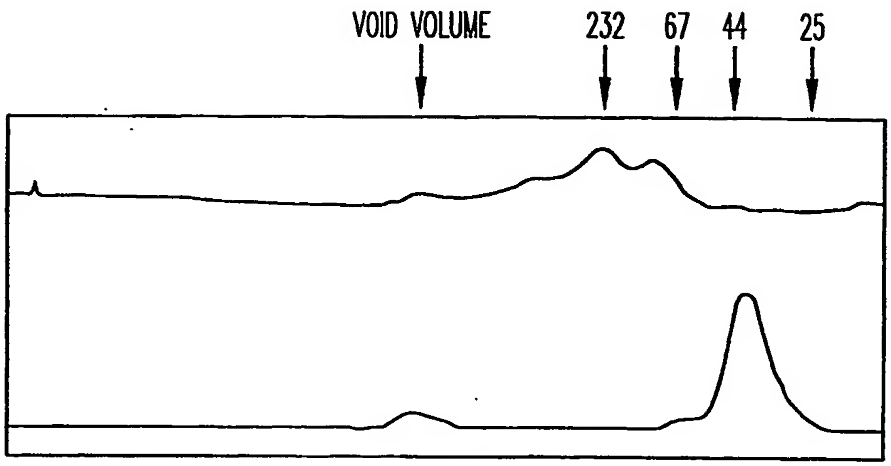


Fig. 10A

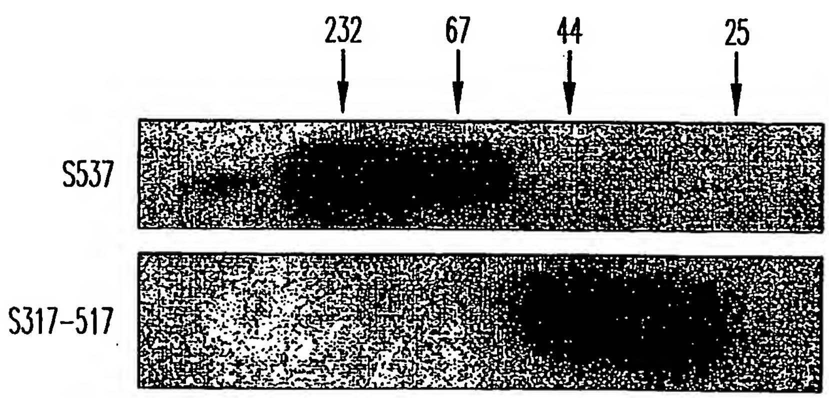


Fig. 10B

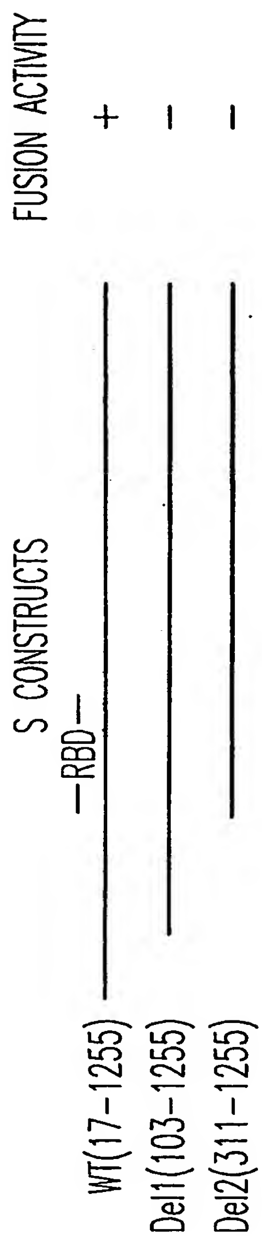

FUSION ACTIVITY
S CONSTRUCTS
RBD
WT(17–1255)
Del1(103–1255)
Del2(311–1255)
+
–
–

Fig. 11A

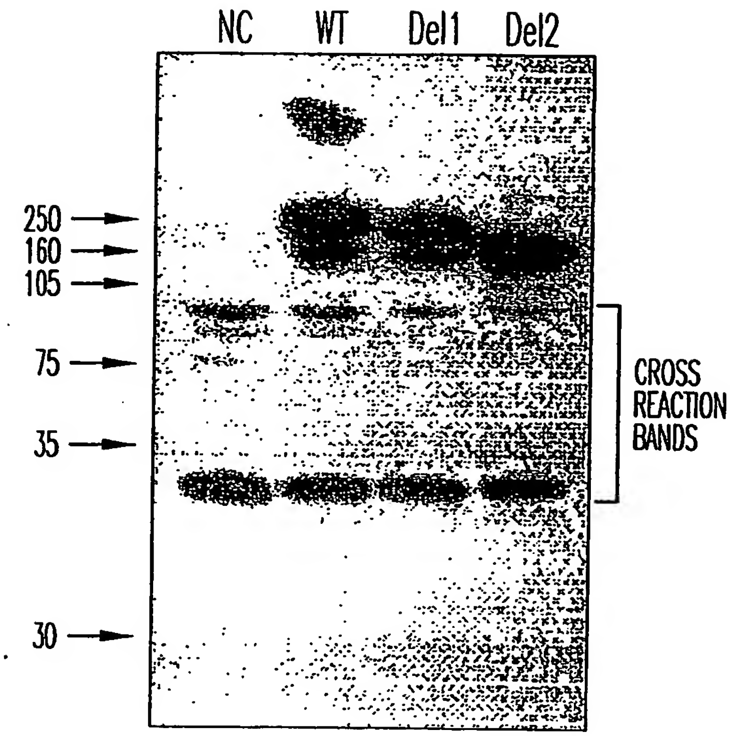


Fig. 11B

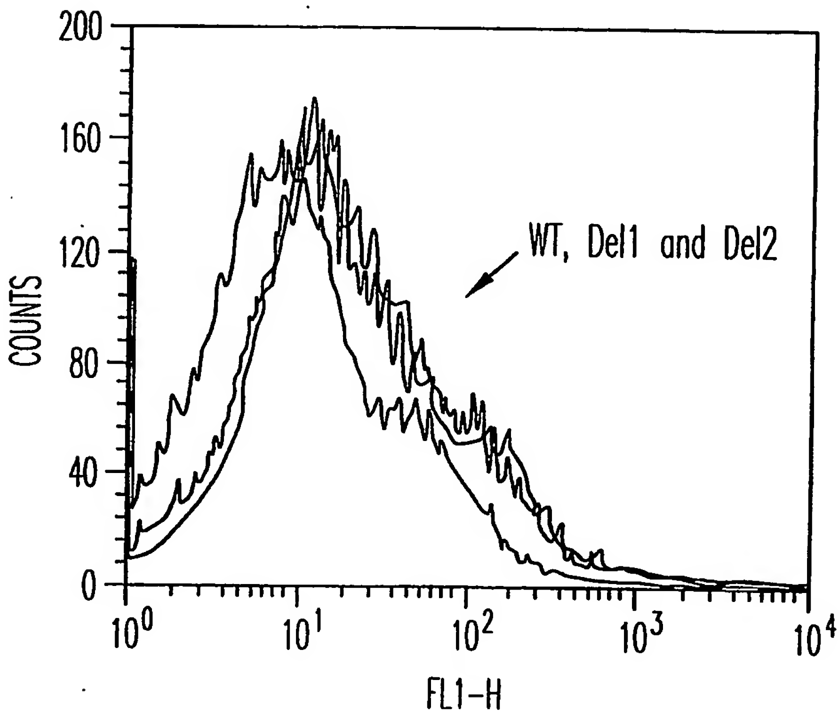


Fig. 11C

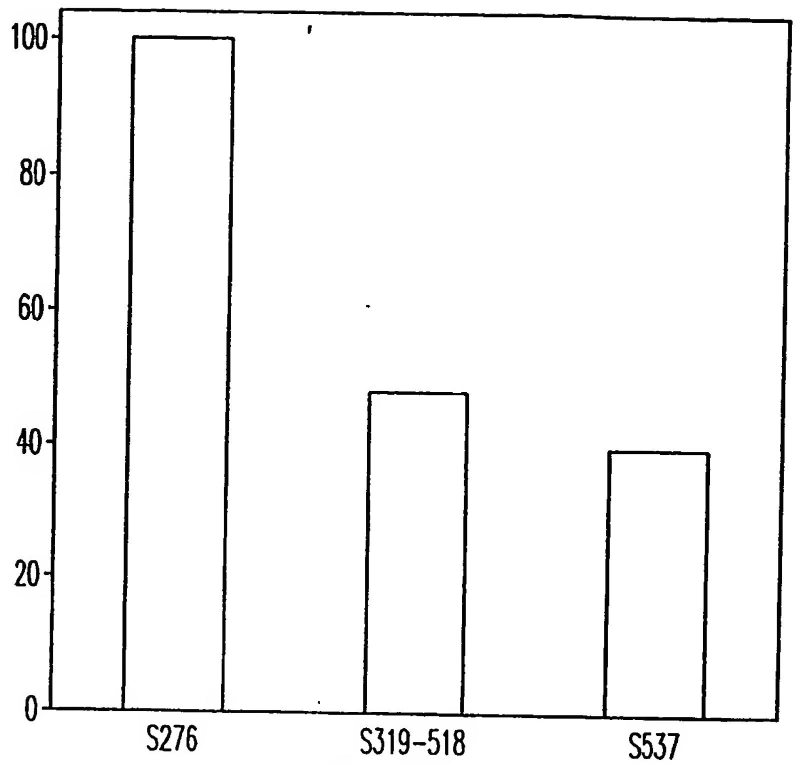


Fig. 12A

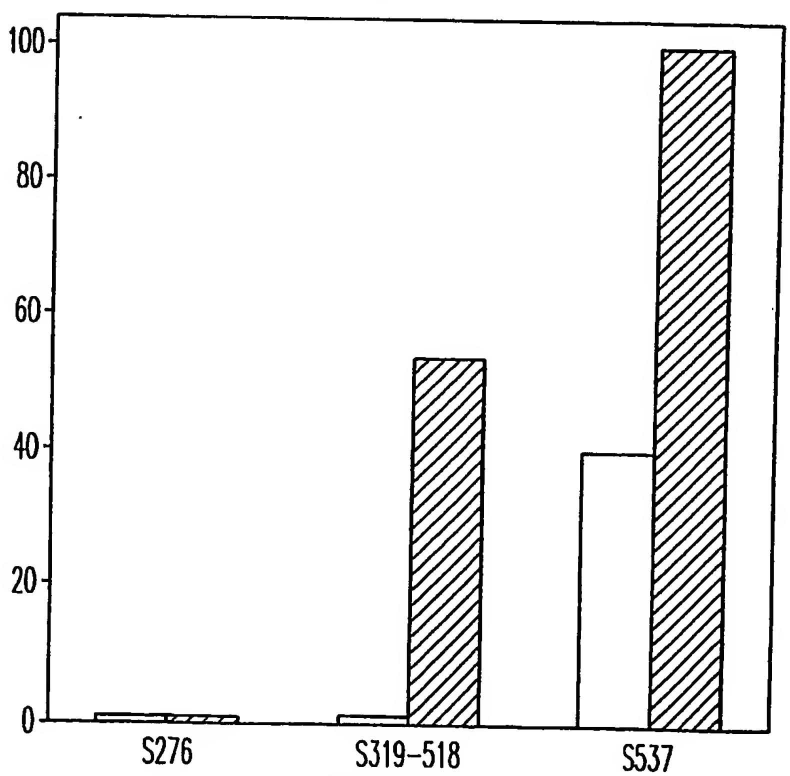


Fig. 12B

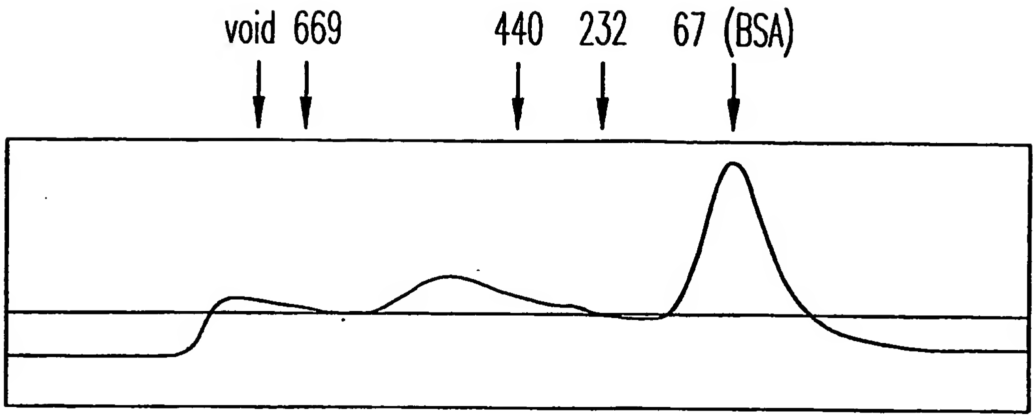


Fig. 13A

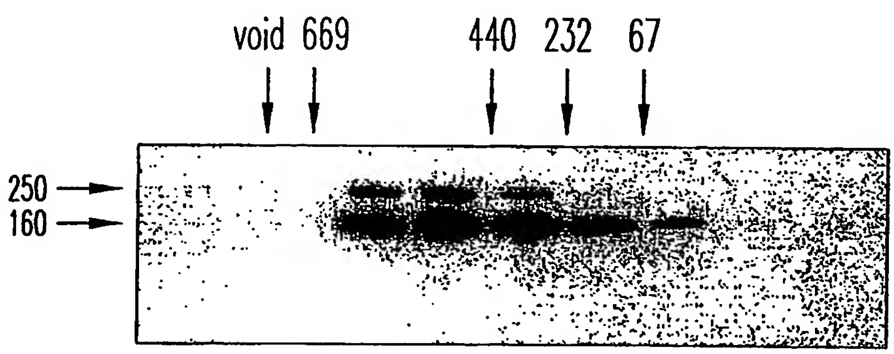


Fig. 13B

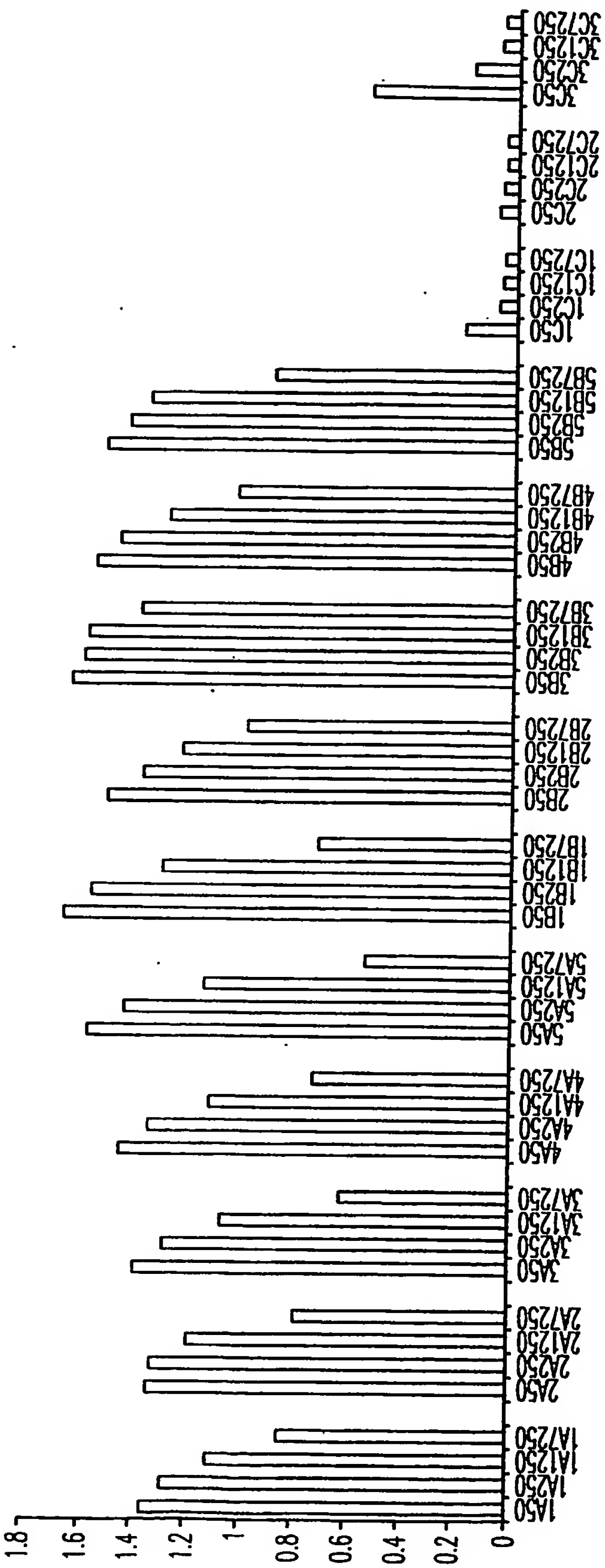
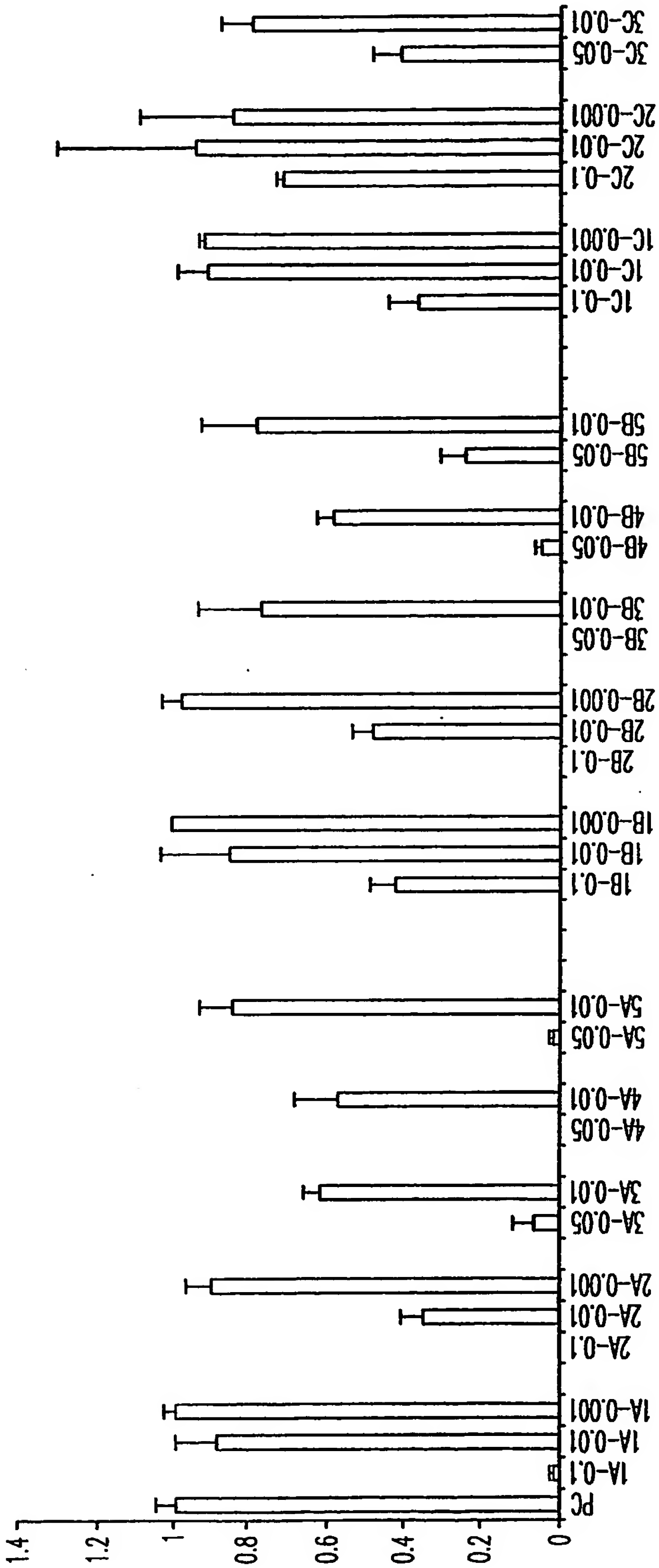

0
0.2
0.4
0.6
0.8
1
1.2
1.4
1.6
1.8
1A50
1A250
1A1250
1A7250
2A50
2A250
2A1250
2A7250
3A50
3A250
3A1250
3A7250
4A50
4A250
4A1250
4A7250
5A50
5A250
5A1250
5A7250
1B50
1B250
1B1250
1B7250
2B50
2B250
2B1250
2B7250
3B50
3B250
3B1250
3B7250
4B50
4B250
4B1250
4B7250
5B50
5B250
5B1250
5B7250
1C50
1C250
1C1250
1C7250
2C50
2C250
2C1250
2C7250
3C50
3C250
3C1250
3C7250

Fig. 14A



1.4
1.2
1
0.8
0.6
0.4
0.2
0
PC
1A-0.1
1A-0.01
1A-0.001
2A-0.1
2A-0.01
2A-0.001
3A-0.05
3A-0.01
4A-0.05
4A-0.01
5A-0.05
5A-0.01
1B-0.1
1B-0.01
1B-0.001
2B-0.1
2B-0.01
2B-0.001
3B-0.05
3B-0.01
4B-0.05
4B-0.01
5B-0.05
5B-0.01
1C-0.1
1C-0.01
1C-0.001
2C-0.1
2C-0.01
2C-0.001
3C-0.05
3C-0.01

Fig. 14B

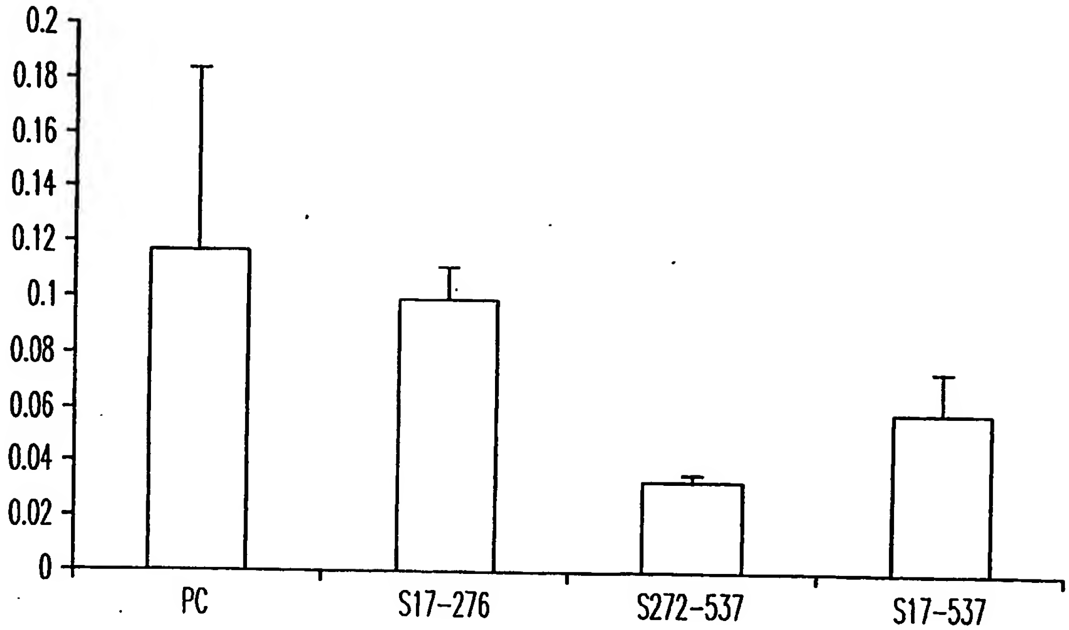

0.2
0.18
0.16
0.14
0.12
0.1
0.08
0.06
0.04
0.02
0
PC
S17-276
S272-537
S17-537

Fig. 15

SEQUENCE LISTING

<110> National Institutes of Health
Dimitrov, Dimiter S.
Xiao, Xiaodong

<120> Soluble Fragments of the SARS-CoV Spike Glycoprotein

<130> 1662.024WO1

<150> US 60/489,166
<151> 2003-07-21

<150> US 60/524,642
<151> 2003-11-25

<160> 65

<170> FastSEQ for Windows Version 4.0

<210> 1
<211> 1255
<212> PRT
<213> SARS coronavirus

<400> 1
Met Phe Ile Phe Leu Leu Phe Leu Thr Leu Thr Ser Gly Ser Asp Leu
1 5 10 15
Asp Arg Cys Thr Thr Phe Asp Asp Val Gln Ala Pro Asn Tyr Thr Gln
20 25 30
His Thr Ser Ser Met Arg Gly Val Tyr Tyr Pro Asp Glu Ile Phe Arg
35 40 45
Ser Asp Thr Leu Tyr Leu Thr Gln Asp Leu Phe Leu Pro Phe Tyr Ser
50 55 60
Asn Val Thr Gly Phe His Thr Ile Asn His Thr Phe Gly Asn Pro Val
65 70 75 80
Ile Pro Phe Lys Asp Gly Ile Tyr Phe Ala Ala Thr Glu Lys Ser Asn
85 90 95
Val Val Arg Gly Trp Val Phe Gly Ser Thr Met Asn Asn Lys Ser Gln
100 105 110
Ser Val Ile Ile Ile Asn Asn Ser Thr Asn Val Val Ile Arg Ala Cys
115 120 125

Asn Phe Glu Leu Cys Asp Asn Pro Phe Phe Ala Val Ser Lys Pro Met
130 135 140

Gly Thr Gln Thr His Thr Met Ile Phe Asp Asn Ala Phe Asn Cys Thr
145 150 155 160

Phe Glu Tyr Ile Ser Asp Ala Phe Ser Leu Asp Val Ser Glu Lys Ser
165 170 175

Gly Asn Phe Lys His Leu Arg Glu Phe Val Phe Lys Asn Lys Asp Gly
180 185 190

Phe Leu Tyr Val Tyr Lys Gly Tyr Gln Pro Ile Asp Val Val Arg Asp
195 200 205

Leu Pro Ser Gly Phe Asn Thr Leu Lys Pro Ile Phe Lys Leu Pro Leu
210 215 220

Gly Ile Asn Ile Thr Asn Phe Arg Ala Ile Leu Thr Ala Phe Ser Pro
225 230 235 240

Ala Gln Asp Ile Trp Gly Thr Ser Ala Ala Ala Tyr Phe Val Gly Tyr
245 250 255

Leu Lys Pro Thr Thr Phe Met Leu Lys Tyr Asp Glu Asn Gly Thr Ile
260 265 270

Thr Asp Ala Val Asp Cys Ser Gln Asn Pro Leu Ala Glu Leu Lys Cys
275 280 285

Ser Val Lys Ser Phe Glu Ile Asp Lys Gly Ile Tyr Gln Thr Ser Asn
290 295 300

Phe Arg Val Val Pro Ser Gly Asp Val Val Arg Phe Pro Asn Ile Thr
305 310 315 320

Asn Leu Cys Pro Phe Gly Glu Val Phe Asn Ala Thr Lys Phe Pro Ser
325 330 335

Val Tyr Ala Trp Glu Arg Lys Lys Ile Ser Asn Cys Val Ala Asp Tyr
340 345 350

Ser Val Leu Tyr Asn Ser Thr Phe Phe Ser Thr Phe Lys Cys Tyr Gly
355 360 365

Val Ser Ala Thr Lys Leu Asn Asp Leu Cys Phe Ser Asn Val Tyr Ala
370 375 380

Asp Ser Phe Val Val Lys Gly Asp Asp Val Arg Gln Ile Ala Pro Gly
385 390 395 400

Gln Thr Gly Val Ile Ala Asp Tyr Asn Tyr Lys Leu Pro Asp Asp Phe
405 410 415

Met Gly Cys Val Leu Ala Trp Asn Thr Arg Asn Ile Asp Ala Thr Ser
420 425 430

Thr Gly Asn Tyr Asn Tyr Lys Tyr Arg Tyr Leu Arg His Gly Lys Leu
435 440 445

Arg Pro Phe Glu Arg Asp Ile Ser Asn Val Pro Phe Ser Pro Asp Gly
450 455 460

Lys Pro Cys Thr Pro Pro Ala Leu Asn Cys Tyr Trp Pro Leu Asn Asp
465 470 475 480
Tyr Gly Phe Tyr Thr Thr Thr Gly Ile Gly Tyr Gln Pro Tyr Arg Val
485 490 495
Val Val Leu Ser Phe Glu Leu Leu Asn Ala Pro Ala Thr Val Cys Gly
500 505 510
Pro Lys Leu Ser Thr Asp Leu Ile Lys Asn Gln Cys Val Asn Phe Asn
515 520 525
Phe Asn Gly Leu Thr Gly Thr Gly Val Leu Thr Pro Ser Ser Lys Arg
530 535 540
Phe Gln Pro Phe Gln Gln Phe Gly Arg Asp Val Ser Asp Phe Thr Asp
545 550 555 560
Ser Val Arg Asp Pro Lys Thr Ser Glu Ile Leu Asp Ile Ser Pro Cys
565 570 575
Ala Phe Gly Gly Val Ser Val Ile Thr Pro Gly Thr Asn Ala Ser Ser
580 585 590
Glu Val Ala Val Leu Tyr Gln Asp Val Asn Cys Thr Asp Val Ser Thr
595 600 605
Ala Ile His Ala Asp Gln Leu Thr Pro Ala Trp Arg Ile Tyr Ser Thr
610 615 620
Gly Asn Asn Val Phe Gln Thr Gln Ala Gly Cys Leu Ile Gly Ala Glu
625 630 635 640
His Val Asp Thr Ser Tyr Glu Cys Asp Ile Pro Ile Gly Ala Gly Ile
645 650 655
Cys Ala Ser Tyr His Thr Val Ser Leu Leu Arg Ser Thr Ser Gln Lys
660 665 670
Ser Ile Val Ala Tyr Thr Met Ser Leu Gly Ala Asp Ser Ser Ile Ala
675 680 685
Tyr Ser Asn Asn Thr Ile Ala Ile Pro Thr Asn Phe Ser Ile Ser Ile
690 695 700
Thr Thr Glu Val Met Pro Val Ser Met Ala Lys Thr Ser Val Asp Cys
705 710 715 720
Asn Met Tyr Ile Cys Gly Asp Ser Thr Glu Cys Ala Asn Leu Leu Leu
725 730 735
Gln Tyr Gly Ser Phe Cys Thr Gln Leu Asn Arg Ala Leu Ser Gly Ile
740 745 750
Ala Ala Glu Gln Asp Arg Asn Thr Arg Glu Val Phe Ala Gln Val Lys
755 760 765
Gln Met Tyr Lys Thr Pro Thr Leu Lys Tyr Phe Gly Gly Phe Asn Phe
770 775 780
Ser Gln Ile Leu Pro Asp Pro Leu Lys Pro Thr Lys Arg Ser Phe Ile
785 790 795 800

Glu Asp Leu Leu Phe Asn Lys Val Thr Leu Ala Asp Ala Gly Phe Met
805 810 815

Lys Gln Tyr Gly Glu Cys Leu Gly Asp Ile Asn Ala Arg Asp Leu Ile
820 825 830

Cys Ala Gln Lys Phe Asn Gly Leu Thr Val Leu Pro Pro Leu Leu Thr
835 840 845

Asp Asp Met Ile Ala Ala Tyr Thr Ala Ala Leu Val Ser Gly Thr Ala
850 855 860

Thr Ala Gly Trp Thr Phe Gly Ala Gly Ala Ala Leu Gln Ile Pro Phe
865 870 875 880

Ala Met Gln Met Ala Tyr Arg Phe Asn Gly Ile Gly Val Thr Gln Asn
885 890 895

Val Leu Tyr Glu Asn Gln Lys Gln Ile Ala Asn Gln Phe Asn Lys Ala
900 905 910

Ile Ser Gln Ile Gln Glu Ser Leu Thr Thr Thr Ser Thr Ala Leu Gly
915 920 925

Lys Leu Gln Asp Val Val Asn Gln Asn Ala Gln Ala Leu Asn Thr Leu
930 935 940

Val Lys Gln Leu Ser Ser Asn Phe Gly Ala Ile Ser Ser Val Leu Asn
945 950 955 960

Asp Ile Leu Ser Arg Leu Asp Lys Val Glu Ala Glu Val Gln Ile Asp
965 970 975

Arg Leu Ile Thr Gly Arg Leu Gln Ser Leu Gln Thr Tyr Val Thr Gln
980 985 990

Gln Leu Ile Arg Ala Ala Glu Ile Arg Ala Ser Ala Asn Leu Ala Ala
995 1000 1005

Thr Lys Met Ser Glu Cys Val Leu Gly Gln Ser Lys Arg Val Asp Phe
1010 1015 1020

Cys Gly Lys Gly Tyr His Leu Met Ser Phe Pro Gln Ala Ala Pro His
1025 1030 1035 1040

Gly Val Val Phe Leu His Val Thr Tyr Val Pro Ser Gln Glu Arg Asn
1045 1050 1055

Phe Thr Thr Ala Pro Ala Ile Cys His Glu Gly Lys Ala Tyr Phe Pro
1060 1065 1070

Arg Glu Gly Val Phe Val Phe Asn Gly Thr Ser Trp Phe Ile Thr Gln
1075 1080 1085

Arg Asn Phe Phe Ser Pro Gln Ile Ile Thr Thr Asp Asn Thr Phe Val
1090 1095 1100

Ser Gly Asn Cys Asp Val Val Ile Gly Ile Ile Asn Asn Thr Val Tyr
1105 1110 1115 1120

Asp Pro Leu Gln Pro Glu Leu Asp Ser Phe Lys Glu Glu Leu Asp Lys
1125 1130 1135

Tyr Phe Lys Asn His Thr Ser Pro Asp Val Asp Leu Gly Asp Ile Ser
1140 1145 1150

Gly Ile Asn Ala Ser Val Val Asn Ile Gln Lys Glu Ile Asp Arg Leu
1155 1160 1165

Asn Glu Val Ala Lys Asn Leu Asn Glu Ser Leu Ile Asp Leu Gln Glu
1170 1175 1180

Leu Gly Lys Tyr Glu Gln Tyr Ile Lys Trp Pro Trp Tyr Val Trp Leu
1185 1190 1195 1200

Gly Phe Ile Ala Gly Leu Ile Ala Ile Val Met Val Thr Ile Leu Leu
1205 1210 1215

Cys Cys Met Thr Ser Cys Cys Ser Cys Leu Lys Gly Ala Cys Ser Cys
1220 1225 1230

Gly Ser Cys Cys Lys Phe Asp Glu Asp Asp Ser Glu Pro Val Leu Lys
1235 1240 1245

Gly Val Lys Leu His Tyr Thr
1250 1255

<210> 2
<211> 3768
<212> DNA
<213> SARS coronavirus

<400> 2
atgtttattt tcttattatt tcttactctc actagtggta gtgaccttga ccggtgcacc 60
acttttgatg atgttcaagc tcctaattac actcaacata cttcatctat gaggggggtt 120
tactatcctg atgaaatttt tagatcagac actctttatt taactcagga tttatttctt 180
ccattttatt ctaatgttac agggtttcat actattaatc atacgtttgg caaccctgtc 240
ataccttta aggatggtat ttattttgct gccacagaga aatcaaatgt tgtccgtggt 300
tgggtttttg gttctaccat gaacaacaag tcacagtcgg tgattattat taacaattct 360
actaatgttg ttatacgagc atgtaacttt gaattgtgtg acaacccttt ctttgctgtt 420
tctaaaccca tgggtacaca gacacatact atgatattcg ataatgcatt taattgcact 480
ttcgagtaca tatctgatgc cttttcgctt gatgtttcag aaaagtcagg taattttaaa 540
cacttacgag agtttgtgtt taaaaataaa gatgggtttc tctatgttta taagggctat 600
caacctatag atgtagttcg tgatctacct tctggtttta acactttgaa acctattttt 660
aagttgcctc ttggtattaa cattacaaat tttagagcca ttcttacagc cttttcacct 720
gctcaagaca tttggggcac gtcagctgca gcctattttg ttggctattt aaagccaact 780
acatttatgc tcaagtatga tgaaaatggt acaatcacag atgctgttga ttgttctcaa 840
aatccacttg ctgaactcaa atgctctgtt aagagctttg agattgacaa aggaatttac 900
cagacctcta atttcagggt tgttccctca ggagatgttg tgagattccc taatattaca 960
aacttgtgtc cttttggaga ggtttttaat gctactaaat tcccttctgt ctatgcatgg 1020
gagagaaaaa aaatttctaa ttgtgttgct gattactctg tgctctacaa ctcaacattt 1080
ttttcaacct ttaagtgcta tggcgtttct gccactaagt tgaatgatct ttgcttctcc 1140

aatgtctatg cagattcttt tgtagtcaag ggagatgatg taagacaaat agcgccagga 1200
caaactggtg ttattgctga ttataattat aaattgccag atgatttcat gggttgtgtc 1260
cttgcttgga atactaggaa cattgatgct acttcaactg gtaattataa ttataaatat 1320
aggtatctta gacatggcaa gcttaggccc tttgagagag acatatctaa tgtgcctttc 1380
tcccctgatg gcaaaccttg caccccacct gctcttaatt gttattggcc attaaatgat 1440
tatggttttt acaccactac tggcattggc taccaacctt acagagttgt agtactttct 1500
tttgaacttt taaatgcacc ggccacggtt tgtggaccaa aattatccac tgaccttatt 1560
aagaaccagt gtgtcaattt taattttaat ggactcactg gtactggtgt gttaactcct 1620
tcttcaaaga gatttcaacc atttcaacaa tttggccgtg atgtttctga tttcactgat 1680
tccgttcgag atcctaaaac atctgaaata ttagacattt caccttgcgc ttttgggggt 1740
gtaagtgtaa ttacacctgg aacaaatgct tcatctgaag ttgctgttct atatcaagat 1800
gttaactgca ctgatgtttc tacagcaatt catgcagatc aactcacacc agcttggcgc 1860
atatattcta ctggaaacaa tgtattccag actcaagcag gctgtcttat aggagctgag 1920
catgtcgaca cttcttatga gtgcgacatt cctattggag ctggcatttg tgctagttac 1980
catacagttt ctttattacg tagtactagc caaaaatcta ttgtggctta tactatgtct 2040
ttaggtgctg atagttcaat tgcttactct aataacacca ttgctatacc tactaacttt 2100
tcaattagca ttactacaga agtaatgcct gtttctatgg ctaaaacctc cgtagattgt 2160
aatatgtaca tctgcggaga ttctactgaa tgtgctaatt tgcttctcca atatggtagc 2220
ttttgcacac aactaaatcg tgcactctca ggtattgctg ctgaacagga tcgcaacaca 2280
cgtgaagtgt tcgctcaagt caaacaaatg tacaaaaccc caactttgaa atattttggt 2340
ggttttaatt tttcacaaat attacctgac cctctaaagc caactaagag gtcttttatt 2400
gaggacttgc tctttaataa ggtgacactc gctgatgctg gcttcatgaa gcaatatggc 2460
gaatgcctag gtgatattaa tgctagagat ctcatttgtg cgcagaagtt caatggactt 2520
acagtgttgc cacctctgct cactgatgat atgattgctg cctacactgc tgctctagtt 2580
agtggtactg ccactgctgg atggacattt ggtgctggcg ctgctcttca aatacctttt 2640
gctatgcaaa tggcatatag gttcaatggc attggagtta cccaaaatgt tctctatgag 2700
aaccaaaaac aaatcgccaa ccaatttaac aaggcgatta gtcaaattca agaatcactt 2760
acaacaacat caactgcatt gggcaagctg caagacgttg ttaaccagaa tgctcaagca 2820
ttaaacacac ttgttaaaca acttagctct aattttggtg caatttcaag tgtgctaaat 2880
gatatccttt cgcgacttga taaagtcgag gcggaggtac aaattgacag gttaattaca 2940
ggcagacttc aaagccttca aacctatgta acacaacaac taatcagggc tgctgaaatc 3000
agggcttctg ctaatcttgc tgctactaaa atgtctgagt gtgttcttgg acaatcaaaa 3060
agagttgact tttgtggaaa gggctaccac cttatgtcct tcccacaagc agccccgcat 3120
ggtgttgtct tcctacatgt cacgtatgtg ccatcccagg agaggaactt caccacagcg 3180
ccagcaattt gtcatgaagg caaagcatac ttccctcgtg aaggtgtttt tgtgtttaat 3240
ggcacttctt ggtttattac acagaggaac ttcttttctc cacaaataat tactacagac 3300
aatacatttg tctcaggaaa ttgtgatgtc gttattggca tcattaacaa cacagtttat 3360
gatcctctgc aacctgagct cgactcattc aaagaagagc tggacaagta cttcaaaaat 3420
catacatcac cagatgttga tcttggcgac atttcaggca ttaacgcttc tgtcgtcaac 3480
attcaaaaag aaattgaccg cctcaatgag gtcgctaaaa atttaaatga atcactcatt 3540
gaccttcaag aattgggaaa atatgagcaa tatattaaat ggccttggta tgtttggctc 3600
ggcttcattg ctggactaat tgccatcgtc atggttacaa tcttgctttg ttgcatgact 3660

agttgttgca gttgcctcaa gggtgcatgc tcttgtggtt cttgctgcaa gtttgatgag 3720
gatgactctg agccagttct caagggtgtc aaattacatt acacataa 3768

<210> 3
<211> 29
<212> DNA
<213> Artificial Sequence

<220>
<223> A synthetic primer

<400> 3
agtcggatcc ggtaggctta tcattagag 29

<210> 4
<211> 20
<212> DNA
<213> Artificial Sequence

<220>
<223> A synthetic primer

<400> 4
ccatcagggg agaaaggcac 20

<210> 5
<211> 20
<212> DNA
<213> Artificial Sequence

<220>
<223> A synthetic primer

<400> 5
gtgcctttct cccctgatgg 20

<210> 6
<211> 19
<212> DNA
<213> Artificial Sequence

<220>
<223> A synthetic primer

<400> 6
gaagagcagc gccagcacc 19

<210> 7
<211> 19
<212> DNA
<213> Artificial Sequence

<220>
<223> A synthetic primer

<400> 7
ggtgctggcg ctgctcttc 19

<210> 8
<211> 28
<212> DNA
<213> Artificial Sequence

<220>
<223> A synthetic primer

<400> 8
actgtctaga gttcgtttat gtgtaatg 28

<210> 9
<211> 29
<212> DNA
<213> Artificial Sequence

<220>
<223> A synthetic primer

<400> 9
agtcggatcc gaccggtgca ccacttttg 29

<210> 10
<211> 28
<212> DNA
<213> Artificial Sequence

<220>
<223> A synthetic primer

<400> 10
agtcgggccc ctgttcagca gcaatacc 28

<210> 11
<211> 28
<212> DNA
<213> Artificial Sequence

<220>
<223> A synthetic primer

<400> 11
actgggatcc gaagtgttcg ctcaagtc 28

<210> 12
<211> 26
<212> DNA
<213> Artificial Sequence

<220>
<223> A synthetic primer

<400> 12
actgtctaga ttgctcatat tttccc 26

<210> 13
<211> 740
<212> PRT
<213> SARS coronavirus

<400> 13
Asp Arg Cys Thr Thr Phe Asp Asp Val Gln Ala Pro Asn Tyr Thr Gln
1 5 10 15

His Thr Ser Ser Met Arg Gly Val Tyr Tyr Pro Asp Glu Ile Phe Arg
20 25 30

Ser Asp Thr Leu Tyr Leu Thr Gln Asp Leu Phe Leu Pro Phe Tyr Ser
35 40 45

Asn Val Thr Gly Phe His Thr Ile Asn His Thr Phe Gly Asn Pro Val
50 55 60

Ile Pro Phe Lys Asp Gly Ile Tyr Phe Ala Ala Thr Glu Lys Ser Asn
65 70 75 80

Val Val Arg Gly Trp Val Phe Gly Ser Thr Met Asn Asn Lys Ser Gln
85 90 95

Ser Val Ile Ile Ile Asn Asn Ser Thr Asn Val Val Ile Arg Ala Cys
100 105 110

Asn Phe Glu Leu Cys Asp Asn Pro Phe Phe Ala Val Ser Lys Pro Met
115 120 125

Gly Thr Gln Thr His Thr Met Ile Phe Asp Asn Ala Phe Asn Cys Thr
130 135 140

Phe Glu Tyr Ile Ser Asp Ala Phe Ser Leu Asp Val Ser Glu Lys Ser
145 150 155 160

Gly Asn Phe Lys His Leu Arg Glu Phe Val Phe Lys Asn Lys Asp Gly
165 170 175

Phe Leu Tyr Val Tyr Lys Gly Tyr Gln Pro Ile Asp Val Val Arg Asp
180 185 190

Leu Pro Ser Gly Phe Asn Thr Leu Lys Pro Ile Phe Lys Leu Pro Leu
195 200 205

Gly Ile Asn Ile Thr Asn Phe Arg Ala Ile Leu Thr Ala Phe Ser Pro
210 215 220

Ala Gln Asp Ile Trp Gly Thr Ser Ala Ala Ala Tyr Phe Val Gly Tyr
225 230 235 240

Leu Lys Pro Thr Thr Phe Met Leu Lys Tyr Asp Glu Asn Gly Thr Ile
245 250 255

Thr Asp Ala Val Asp Cys Ser Gln Asn Pro Leu Ala Glu Leu Lys Cys
260 265 270

Ser Val Lys Ser Phe Glu Ile Asp Lys Gly Ile Tyr Gln Thr Ser Asn
275 280 285

Phe Arg Val Val Pro Ser Gly Asp Val Val Arg Phe Pro Asn Ile Thr
290 295 300

Asn Leu Cys Pro Phe Gly Glu Val Phe Asn Ala Thr Lys Phe Pro Ser
305 310 315 320

Val Tyr Ala Trp Glu Arg Lys Lys Ile Ser Asn Cys Val Ala Asp Tyr
325 330 335

Ser Val Leu Tyr Asn Ser Thr Phe Phe Ser Thr Phe Lys Cys Tyr Gly
340 345 350

Val Ser Ala Thr Lys Leu Asn Asp Leu Cys Phe Ser Asn Val Tyr Ala
355 360 365

Asp Ser Phe Val Val Lys Gly Asp Asp Val Arg Gln Ile Ala Pro Gly
370 375 380

Gln Thr Gly Val Ile Ala Asp Tyr Asn Tyr Lys Leu Pro Asp Asp Phe
385 390 395 400

Met Gly Cys Val Leu Ala Trp Asn Thr Arg Asn Ile Asp Ala Thr Ser
405 410 415

Thr Gly Asn Tyr Asn Tyr Lys Tyr Arg Tyr Leu Arg His Gly Lys Leu
420 425 430

Arg Pro Phe Glu Arg Asp Ile Ser Asn Val Pro Phe Ser Pro Asp Gly
435 440 445

Lys Pro Cys Thr Pro Pro Ala Leu Asn Cys Tyr Trp Pro Leu Asn Asp
450 455 460

Tyr Gly Phe Tyr Thr Thr Thr Gly Ile Gly Tyr Gln Pro Tyr Arg Val
465 470 475 480

Val Val Leu Ser Phe Glu Leu Leu Asn Ala Pro Ala Thr Val Cys Gly
485 490 495

Pro Lys Leu Ser Thr Asp Leu Ile Lys Asn Gln Cys Val Asn Phe Asn
500 505 510

Phe Asn Gly Leu Thr Gly Thr Gly Val Leu Thr Pro Ser Ser Lys Arg
515 520 525

Phe Gln Pro Phe Gln Gln Phe Gly Arg Asp Val Ser Asp Phe Thr Asp
530 535 540

Ser Val Arg Asp Pro Lys Thr Ser Glu Ile Leu Asp Ile Ser Pro Cys
545 550 555 560

Ala Phe Gly Gly Val Ser Val Ile Thr Pro Gly Thr Asn Ala Ser Ser
565 570 575

Glu Val Ala Val Leu Tyr Gln Asp Val Asn Cys Thr Asp Val Ser Thr
580 585 590

Ala Ile His Ala Asp Gln Leu Thr Pro Ala Trp Arg Ile Tyr Ser Thr
595 600 605

Gly Asn Asn Val Phe Gln Thr Gln Ala Gly Cys Leu Ile Gly Ala Glu
610 615 620

His Val Asp Thr Ser Tyr Glu Cys Asp Ile Pro Ile Gly Ala Gly Ile
625 630 635 640

Cys Ala Ser Tyr His Thr Val Ser Leu Leu Arg Ser Thr Ser Gln Lys
645 650 655

Ser Ile Val Ala Tyr Thr Met Ser Leu Gly Ala Asp Ser Ser Ile Ala
660 665 670

Tyr Ser Asn Asn Thr Ile Ala Ile Pro Thr Asn Phe Ser Ile Ser Ile
675 680 685

Thr Thr Glu Val Met Pro Val Ser Met Ala Lys Thr Ser Val Asp Cys
690 695 700
Asn Met Tyr Ile Cys Gly Asp Ser Thr Glu Cys Ala Asn Leu Leu Leu
705 710 715 720
Gln Tyr Gly Ser Phe Cys Thr Gln Leu Asn Arg Ala Leu Ser Gly Ile
725 730 735
Ala Ala Glu Gln
740

<210> 14
<211> 429
<212> PRT
<213> SARS coronavirus

<400> 14
Glu Val Phe Ala Gln Val Lys Gln Met Tyr Lys Thr Pro Thr Leu Lys
1 5 10 15
Tyr Phe Gly Gly Phe Asn Phe Ser Gln Ile Leu Pro Asp Pro Leu Lys
20 25 30
Pro Thr Lys Arg Ser Phe Ile Glu Asp Leu Leu Phe Asn Lys Val Thr
35 40 45
Leu Ala Asp Ala Gly Phe Met Lys Gln Tyr Gly Glu Cys Leu Gly Asp
50 55 60
Ile Asn Ala Arg Asp Leu Ile Cys Ala Gln Lys Phe Asn Gly Leu Thr
65 70 75 80
Val Leu Pro Pro Leu Leu Thr Asp Asp Met Ile Ala Ala Tyr Thr Ala
85 90 95
Ala Leu Val Ser Gly Thr Ala Thr Ala Gly Trp Thr Phe Gly Ala Gly
100 105 110
Ala Ala Leu Gln Ile Pro Phe Ala Met Gln Met Ala Tyr Arg Phe Asn
115 120 125
Gly Ile Gly Val Thr Gln Asn Val Leu Tyr Glu Asn Gln Lys Gln Ile
130 135 140
Ala Asn Gln Phe Asn Lys Ala Ile Ser Gln Ile Gln Glu Ser Leu Thr
145 150 155 160
Thr Thr Ser Thr Ala Leu Gly Lys Leu Gln Asp Val Val Asn Gln Asn
165 170 175
Ala Gln Ala Leu Asn Thr Leu Val Lys Gln Leu Ser Ser Asn Phe Gly
180 185 190
Ala Ile Ser Ser Val Leu Asn Asp Ile Leu Ser Arg Leu Asp Lys Val
195 200 205

Glu Ala Glu Val Gln Ile Asp Arg Leu Ile Thr Gly Arg Leu Gln Ser
210 215 220
Leu Gln Thr Tyr Val Thr Gln Gln Leu Ile Arg Ala Ala Glu Ile Arg
225 230 235 240
Ala Ser Ala Asn Leu Ala Ala Thr Lys Met Ser Glu Cys Val Leu Gly
245 250 255
Gln Ser Lys Arg Val Asp Phe Cys Gly Lys Gly Tyr His Leu Met Ser
260 265 270
Phe Pro Gln Ala Ala Pro His Gly Val Val Phe Leu His Val Thr Tyr
275 280 285
Val Pro Ser Gln Glu Arg Asn Phe Thr Thr Ala Pro Ala Ile Cys His
290 295 300
Glu Gly Lys Ala Tyr Phe Pro Arg Glu Gly Val Phe Val Phe Asn Gly
305 310 315 320
Thr Ser Trp Phe Ile Thr Gln Arg Asn Phe Phe Ser Pro Gln Ile Ile
325 330 335
Thr Thr Asp Asn Thr Phe Val Ser Gly Asn Cys Asp Val Val Ile Gly
340 345 350
Ile Ile Asn Asn Thr Val Tyr Asp Pro Leu Gln Pro Glu Leu Asp Ser
355 360 365
Phe Lys Glu Glu Leu Asp Lys Tyr Phe Lys Asn His Thr Ser Pro Asp
370 375 380
Val Asp Leu Gly Asp Ile Ser Gly Ile Asn Ala Ser Val Val Asn Ile
385 390 395 400
Gln Lys Glu Ile Asp Arg Leu Asn Glu Val Ala Lys Asn Leu Asn Glu
405 410 415
Ser Leu Ile Asp Leu Gln Glu Leu Gly Lys Tyr Glu Gln
420 425

<210> 15
<211> 1170
<212> PRT
<213> SARS coronavirus

<400> 15
Asp Arg Cys Thr Thr Phe Asp Asp Val Gln Ala Pro Asn Tyr Thr Gln
1 5 10 15
His Thr Ser Ser Met Arg Gly Val Tyr Tyr Pro Asp Glu Ile Phe Arg
20 25 30
Ser Asp Thr Leu Tyr Leu Thr Gln Asp Leu Phe Leu Pro Phe Tyr Ser
35 40 45

Asn Val Thr Gly Phe His Thr Ile Asn His Thr Phe Gly Asn Pro Val
50 55 60

Ile Pro Phe Lys Asp Gly Ile Tyr Phe Ala Ala Thr Glu Lys Ser Asn
65 70 75 80

Val Val Arg Gly Trp Val Phe Gly Ser Thr Met Asn Asn Lys Ser Gln
85 90 95

Ser Val Ile Ile Ile Asn Asn Ser Thr Asn Val Val Ile Arg Ala Cys
100 105 110

Asn Phe Glu Leu Cys Asp Asn Pro Phe Phe Ala Val Ser Lys Pro Met
115 120 125

Gly Thr Gln Thr His Thr Met Ile Phe Asp Asn Ala Phe Asn Cys Thr
130 135 140

Phe Glu Tyr Ile Ser Asp Ala Phe Ser Leu Asp Val Ser Glu Lys Ser
145 150 155 160

Gly Asn Phe Lys His Leu Arg Glu Phe Val Phe Lys Asn Lys Asp Gly
165 170 175

Phe Leu Tyr Val Tyr Lys Gly Tyr Gln Pro Ile Asp Val Val Arg Asp
180 185 190

Leu Pro Ser Gly Phe Asn Thr Leu Lys Pro Ile Phe Lys Leu Pro Leu
195 200 205

Gly Ile Asn Ile Thr Asn Phe Arg Ala Ile Leu Thr Ala Phe Ser Pro
210 215 220

Ala Gln Asp Ile Trp Gly Thr Ser Ala Ala Ala Tyr Phe Val Gly Tyr
225 230 235 240

Leu Lys Pro Thr Thr Phe Met Leu Lys Tyr Asp Glu Asn Gly Thr Ile
245 250 255

Thr Asp Ala Val Asp Cys Ser Gln Asn Pro Leu Ala Glu Leu Lys Cys
260 265 270

Ser Val Lys Ser Phe Glu Ile Asp Lys Gly Ile Tyr Gln Thr Ser Asn
275 280 285

Phe Arg Val Val Pro Ser Gly Asp Val Val Arg Phe Pro Asn Ile Thr
290 295 300

Asn Leu Cys Pro Phe Gly Glu Val Phe Asn Ala Thr Lys Phe Pro Ser
305 310 315 320

Val Tyr Ala Trp Glu Arg Lys Lys Ile Ser Asn Cys Val Ala Asp Tyr
325 330 335

Ser Val Leu Tyr Asn Ser Thr Phe Phe Ser Thr Phe Lys Cys Tyr Gly
340 345 350

Val Ser Ala Thr Lys Leu Asn Asp Leu Cys Phe Ser Asn Val Tyr Ala
355 360 365

Asp Ser Phe Val Val Lys Gly Asp Asp Val Arg Gln Ile Ala Pro Gly
370 375 380

Gln Thr Gly Val Ile Ala Asp Tyr Asn Tyr Lys Leu Pro Asp Asp Phe
385 390 395 400
Met Gly Cys Val Leu Ala Trp Asn Thr Arg Asn Ile Asp Ala Thr Ser
405 410 415
Thr Gly Asn Tyr Asn Tyr Lys Tyr Arg Tyr Leu Arg His Gly Lys Leu
420 425 430
Arg Pro Phe Glu Arg Asp Ile Ser Asn Val Pro Phe Ser Pro Asp Gly
435 440 445
Lys Pro Cys Thr Pro Pro Ala Leu Asn Cys Tyr Trp Pro Leu Asn Asp
450 455 460
Tyr Gly Phe Tyr Thr Thr Thr Gly Ile Gly Tyr Gln Pro Tyr Arg Val
465 470 475 480
Val Val Leu Ser Phe Glu Leu Leu Asn Ala Pro Ala Thr Val Cys Gly
485 490 495
Pro Lys Leu Ser Thr Asp Leu Ile Lys Asn Gln Cys Val Asn Phe Asn
500 505 510
Phe Asn Gly Leu Thr Gly Thr Gly Val Leu Thr Pro Ser Ser Lys Arg
515 520 525
Phe Gln Pro Phe Gln Gln Phe Gly Arg Asp Val Ser Asp Phe Thr Asp
530 535 540
Ser Val Arg Asp Pro Lys Thr Ser Glu Ile Leu Asp Ile Ser Pro Cys
545 550 555 560
Ala Phe Gly Gly Val Ser Val Ile Thr Pro Gly Thr Asn Ala Ser Ser
565 570 575
Glu Val Ala Val Leu Tyr Gln Asp Val Asn Cys Thr Asp Val Ser Thr
580 585 590
Ala Ile His Ala Asp Gln Leu Thr Pro Ala Trp Arg Ile Tyr Ser Thr
595 600 605
Gly Asn Asn Val Phe Gln Thr Gln Ala Gly Cys Leu Ile Gly Ala Glu
610 615 620
His Val Asp Thr Ser Tyr Glu Cys Asp Ile Pro Ile Gly Ala Gly Ile
625 630 635 640
Cys Ala Ser Tyr His Thr Val Ser Leu Leu Arg Ser Thr Ser Gln Lys
645 650 655
Ser Ile Val Ala Tyr Thr Met Ser Leu Gly Ala Asp Ser Ser Ile Ala
660 665 670
Tyr Ser Asn Asn Thr Ile Ala Ile Pro Thr Asn Phe Ser Ile Ser Ile
675 680 685
Thr Thr Glu Val Met Pro Val Ser Met Ala Lys Thr Ser Val Asp Cys
690 695 700
Asn Met Tyr Ile Cys Gly Asp Ser Thr Glu Cys Ala Asn Leu Leu Leu
705 710 715 720

Gln Tyr Gly Ser Phe Cys Thr Gln Leu Asn Arg Ala Leu Ser Gly Ile
725 730 735

Ala Ala Glu Gln Asp Glu Val Phe Ala Gln Val Lys Gln Met Tyr Lys
740 745 750

Thr Pro Thr Leu Lys Tyr Phe Gly Gly Phe Asn Phe Ser Gln Ile Leu
755 760 765

Pro Asp Pro Leu Lys Pro Thr Lys Arg Ser Phe Ile Glu Asp Leu Leu
770 775 780

Phe Asn Lys Val Thr Leu Ala Asp Ala Gly Phe Met Lys Gln Tyr Gly
785 790 795 800

Glu Cys Leu Gly Asp Ile Asn Ala Arg Asp Leu Ile Cys Ala Gln Lys
805 810 815

Phe Asn Gly Leu Thr Val Leu Pro Pro Leu Leu Thr Asp Asp Met Ile
820 825 830

Ala Ala Tyr Thr Ala Ala Leu Val Ser Gly Thr Ala Thr Ala Gly Trp
835 840 845

Thr Phe Gly Ala Gly Ala Ala Leu Gln Ile Pro Phe Ala Met Gln Met
850 855 860

Ala Tyr Arg Phe Asn Gly Ile Gly Val Thr Gln Asn Val Leu Tyr Glu
865 870 875 880

Asn Gln Lys Gln Ile Ala Asn Gln Phe Asn Lys Ala Ile Ser Gln Ile
885 890 895

Gln Glu Ser Leu Thr Thr Thr Ser Thr Ala Leu Gly Lys Leu Gln Asp
900 905 910

Val Val Asn Gln Asn Ala Gln Ala Leu Asn Thr Leu Val Lys Gln Leu
915 920 925

Ser Ser Asn Phe Gly Ala Ile Ser Ser Val Leu Asn Asp Ile Leu Ser
930 935 940

Arg Leu Asp Lys Val Glu Ala Glu Val Gln Ile Asp Arg Leu Ile Thr
945 950 955 960

Gly Arg Leu Gln Ser Leu Gln Thr Tyr Val Thr Gln Gln Leu Ile Arg
965 970 975

Ala Ala Glu Ile Arg Ala Ser Ala Asn Leu Ala Ala Thr Lys Met Ser
980 985 990

Glu Cys Val Leu Gly Gln Ser Lys Arg Val Asp Phe Cys Gly Lys Gly
995 1000 1005

Tyr His Leu Met Ser Phe Pro Gln Ala Ala Pro His Gly Val Val Phe
1010 1015 1020

Leu His Val Thr Tyr Val Pro Ser Gln Glu Arg Asn Phe Thr Thr Ala
1025 1030 1035 1040

Pro Ala Ile Cys His Glu Gly Lys Ala Tyr Phe Pro Arg Glu Gly Val
1045 1050 1055

Phe Val Phe Asn Gly Thr Ser Trp Phe Ile Thr Gln Arg Asn Phe Phe
1060 1065 1070

Ser Pro Gln Ile Ile Thr Thr Asp Asn Thr Phe Val Ser Gly Asn Cys
1075 1080 1085

Asp Val Val Ile Gly Ile Ile Asn Asn Thr Val Tyr Asp Pro Leu Gln
1090 1095 1100

Pro Glu Leu Asp Ser Phe Lys Glu Glu Leu Asp Lys Tyr Phe Lys Asn
1105 1110 1115 1120

His Thr Ser Pro Asp Val Asp Leu Gly Asp Ile Ser Gly Ile Asn Ala
1125 1130 1135

Ser Val Val Asn Ile Gln Lys Glu Ile Asp Arg Leu Asn Glu Val Ala
1140 1145 1150

Lys Asn Leu Asn Glu Ser Leu Ile Asp Leu Gln Glu Leu Gly Lys Tyr
1155 1160 1165

Glu Gln
1170

<210> 16
<211> 21
<212> PRT
<213> Artificial Sequence

<220>
<223> A synthetic k chain leader sequence

<400> 16
Met Glu Thr Asp Thr Leu Leu Leu Trp Val Leu Leu Leu Trp Val Pro
1 5 10 15

Gly Ser Thr Gly Asp
20

<210> 17
<211> 10
<212> PRT
<213> Artificial Sequence

<220>
<223> A synthetic myc epitope

<400> 17
Glu Gln Lys Leu Ile Ser Glu Glu Asp Leu
1 5 10

<210> 18
<211> 6
<212> PRT
<213> Artificial Sequence

<220>
<223> A synthetic histidine tag

<400> 18
His His His His His His
1 5

<210> 19
<211> 24
<212> DNA
<213> Artificial Sequence

<220>
<223> A synthetic primer

<400> 19
ctagctcgag caacagcatc tgtg 24

<210> 20
<211> 100
<212> PRT
<213> SARS coronavirus

<400> 20
Met Phe Ile Phe Leu Leu Phe Leu Thr Leu Thr Ser Gly Ser Asp Leu
1 5 10 15
Asp Arg Cys Thr Thr Phe Asp Asp Val Gln Ala Pro Asn Tyr Thr Gln
20 25 30
His Thr Ser Ser Met Arg Gly Val Tyr Tyr Pro Asp Glu Ile Phe Arg
35 40 45
Ser Asp Thr Leu Tyr Leu Thr Gln Asp Leu Phe Leu Pro Phe Tyr Ser
50 55 60
Asn Val Thr Gly Phe His Thr Ile Asn His Thr Phe Gly Asn Pro Val
65 70 75 80
Ile Pro Phe Lys Asp Gly Ile Tyr Phe Ala Ala Thr Glu Lys Ser Asn
85 90 95

Val Val Arg Gly
100

<210> 21
<211> 100
<212> PRT
<213> SARS coronavirus

<400> 21
Trp Val Phe Gly Ser Thr Met Asn Asn Lys Ser Gln Ser Val Ile Ile
1 5 10 15
Ile Asn Asn Ser Thr Asn Val Val Ile Arg Ala Cys Asn Phe Glu Leu
20 25 30
Cys Asp Asn Pro Phe Phe Ala Val Ser Lys Pro Met Gly Thr Gln Thr
35 40 45
His Thr Met Ile Phe Asp Asn Ala Phe Asn Cys Thr Phe Glu Tyr Ile
50 55 60
Ser Asp Ala Phe Ser Leu Asp Val Ser Glu Lys Ser Gly Asn Phe Lys
65 70 75 80
His Leu Arg Glu Phe Val Phe Lys Asn Lys Asp Gly Phe Leu Tyr Val
85 90 95
Tyr Lys Gly Tyr
100

<210> 22
<211> 100
<212> PRT
<213> SARS coronavirus

<400> 22
Gln Pro Ile Asp Val Val Arg Asp Leu Pro Ser Gly Phe Asn Thr Leu
1 5 10 15
Lys Pro Ile Phe Lys Leu Pro Leu Gly Ile Asn Ile Thr Asn Phe Arg
20 25 30
Ala Ile Leu Thr Ala Phe Ser Pro Ala Gln Asp Ile Trp Gly Thr Ser
35 40 45
Ala Ala Ala Tyr Phe Val Gly Tyr Leu Lys Pro Thr Thr Phe Met Leu
50 55 60
Lys Tyr Asp Glu Asn Gly Thr Ile Thr Asp Ala Val Asp Cys Ser Gln
65 70 75 80
Asn Pro Leu Ala Glu Leu Lys Cys Ser Val Lys Ser Phe Glu Ile Asp
85 90 95

Lys Gly Ile Tyr
100

<210> 23
<211> 100
<212> PRT
<213> SARS coronavirus

<400> 23
Gln Thr Ser Asn Phe Arg Val Val Pro Ser Gly Asp Val Val Arg Phe
1 5 10 15
Pro Asn Ile Thr Asn Leu Cys Pro Phe Gly Glu Val Phe Asn Ala Thr
20 25 30
Lys Phe Pro Ser Val Tyr Ala Trp Glu Arg Lys Lys Ile Ser Asn Cys
35 40 45
Val Ala Asp Tyr Ser Val Leu Tyr Asn Ser Thr Phe Phe Ser Thr Phe
50 55 60
Lys Cys Tyr Gly Val Ser Ala Thr Lys Leu Asn Asp Leu Cys Phe Ser
65 70 75 80
Asn Val Tyr Ala Asp Ser Phe Val Val Lys Gly Asp Asp Val Arg Gln
85 90 95
Ile Ala Pro Gly
100

<210> 24
<211> 100
<212> PRT
<213> SARS coronavirus

<400> 24
Gln Thr Gly Val Ile Ala Asp Tyr Asn Tyr Lys Leu Pro Asp Asp Phe
1 5 10 15
Met Gly Cys Val Leu Ala Trp Asn Thr Arg Asn Ile Asp Ala Thr Ser
20 25 30
Thr Gly Asn Tyr Asn Tyr Lys Tyr Arg Tyr Leu Arg His Gly Lys Leu
35 40 45
Arg Pro Phe Glu Arg Asp Ile Ser Asn Val Pro Phe Ser Pro Asp Gly
50 55 60
Lys Pro Cys Thr Pro Pro Ala Leu Asn Cys Tyr Trp Pro Leu Asn Asp
65 70 75 80
Tyr Gly Phe Tyr Thr Thr Thr Gly Ile Gly Tyr Gln Pro Tyr Arg Val
85 90 95

Val Val Leu Ser
100

<210> 25
<211> 100
<212> PRT
<213> SARS coronavirus

<400> 25
Phe Glu Leu Leu Asn Ala Pro Ala Thr Val Cys Gly Pro Lys Leu Ser
1 5 10 15
Thr Asp Leu Ile Lys Asn Gln Cys Val Asn Phe Asn Phe Asn Gly Leu
20 25 30
Thr Gly Thr Gly Val Leu Thr Pro Ser Ser Lys Arg Phe Gln Pro Phe
35 40 45
Gln Gln Phe Gly Arg Asp Val Ser Asp Phe Thr Asp Ser Val Arg Asp
50 55 60
Pro Lys Thr Ser Glu Ile Leu Asp Ile Ser Pro Cys Ala Phe Gly Gly
65 70 75 80
Val Ser Val Ile Thr Pro Gly Thr Asn Ala Ser Ser Glu Val Ala Val
85 90 95
Leu Tyr Gln Asp
100

<210> 26
<211> 100
<212> PRT
<213> SARS coronavirus

<400> 26
Val Asn Cys Thr Asp Val Ser Thr Ala Ile His Ala Asp Gln Leu Thr
1 5 10 15
Pro Ala Trp Arg Ile Tyr Ser Thr Gly Asn Asn Val Phe Gln Thr Gln
20 25 30
Ala Gly Cys Leu Ile Gly Ala Glu His Val Asp Thr Ser Tyr Glu Cys
35 40 45
Asp Ile Pro Ile Gly Ala Gly Ile Cys Ala Ser Tyr His Thr Val Ser
50 55 60
Leu Leu Arg Ser Thr Ser Gln Lys Ser Ile Val Ala Tyr Thr Met Ser
65 70 75 80
Leu Gly Ala Asp Ser Ser Ile Ala Tyr Ser Asn Asn Thr Ile Ala Ile
85 90 95

Pro Thr Asn Phe
100

<210> 27
<211> 100
<212> PRT
<213> SARS coronavirus

<400> 27
Ser Ile Ser Ile Thr Thr Glu Val Met Pro Val Ser Met Ala Lys Thr
1 5 10 15
Ser Val Asp Cys Asn Met Tyr Ile Cys Gly Asp Ser Thr Glu Cys Ala
20 25 30
Asn Leu Leu Leu Gln Tyr Gly Ser Phe Cys Thr Gln Leu Asn Arg Ala
35 40 45
Leu Ser Gly Ile Ala Ala Glu Gln Asp Arg Asn Thr Arg Glu Val Phe
50 55 60
Ala Gln Val Lys Gln Met Tyr Lys Thr Pro Thr Leu Lys Tyr Phe Gly
65 70 75 80
Gly Phe Asn Phe Ser Gln Ile Leu Pro Asp Pro Leu Lys Pro Thr Lys
85 90 95
Arg Ser Phe Ile
100

<210> 28
<211> 100
<212> PRT
<213> SARS coronavirus

<400> 28
Glu Asp Leu Leu Phe Asn Lys Val Thr Leu Ala Asp Ala Gly Phe Met
1 5 10 15
Lys Gln Tyr Gly Glu Cys Leu Gly Asp Ile Asn Ala Arg Asp Leu Ile
20 25 30
Cys Ala Gln Lys Phe Asn Gly Leu Thr Val Leu Pro Pro Leu Leu Thr
35 40 45
Asp Asp Met Ile Ala Ala Tyr Thr Ala Ala Leu Val Ser Gly Thr Ala
50 55 60
Thr Ala Gly Trp Thr Phe Gly Ala Gly Ala Ala Leu Gln Ile Pro Phe
65 70 75 80
Ala Met Gln Met Ala Tyr Arg Phe Asn Gly Ile Gly Val Thr Gln Asn
85 90 95

Val Leu Tyr Glu
100

<210> 29
<211> 100
<212> PRT
<213> SARS coronavirus

<400> 29
Asn Gln Lys Gln Ile Ala Asn Gln Phe Asn Lys Ala Ile Ser Gln Ile
1 5 10 15
Gln Glu Ser Leu Thr Thr Thr Ser Thr Ala Leu Gly Lys Leu Gln Asp
20 25 30
Val Val Asn Gln Asn Ala Gln Ala Leu Asn Thr Leu Val Lys Gln Leu
35 40 45
Ser Ser Asn Phe Gly Ala Ile Ser Ser Val Leu Asn Asp Ile Leu Ser
50 55 60
Arg Leu Asp Lys Val Glu Ala Glu Val Gln Ile Asp Arg Leu Ile Thr
65 70 75 80
Gly Arg Leu Gln Ser Leu Gln Thr Tyr Val Thr Gln Gln Leu Ile Arg
85 90 95
Ala Ala Glu Ile
100

<210> 30
<211> 100
<212> PRT
<213> SARS coronavirus

<400> 30
Arg Ala Ser Ala Asn Leu Ala Ala Thr Lys Met Ser Glu Cys Val Leu
1 5 10 15
Gly Gln Ser Lys Arg Val Asp Phe Cys Gly Lys Gly Tyr His Leu Met
20 25 30
Ser Phe Pro Gln Ala Ala Pro His Gly Val Val Phe Leu His Val Thr
35 40 45
Tyr Val Pro Ser Gln Glu Arg Asn Phe Thr Thr Ala Pro Ala Ile Cys
50 55 60
His Glu Gly Lys Ala Tyr Phe Pro Arg Glu Gly Val Phe Val Phe Asn
65 70 75 80
Gly Thr Ser Trp Phe Ile Thr Gln Arg Asn Phe Phe Ser Pro Gln Ile
85 90 95

Ile Thr Thr Asp
100

<210> 31
<211> 90
<212> PRT
<213> SARS coronavirus

<400> 31
Asn Thr Phe Val Ser Gly Asn Cys Asp Val Val Ile Gly Ile Ile Asn
1 5 10 15
Asn Thr Val Tyr Asp Pro Leu Gln Pro Glu Leu Asp Ser Phe Lys Glu
20 25 30
Glu Leu Asp Lys Tyr Phe Lys Asn His Thr Ser Pro Asp Val Asp Leu
35 40 45
Gly Asp Ile Ser Gly Ile Asn Ala Ser Val Val Asn Ile Gln Lys Glu
50 55 60
Ile Asp Arg Leu Asn Glu Val Ala Lys Asn Leu Asn Glu Ser Leu Ile
65 70 75 80
Asp Leu Gln Glu Leu Gly Lys Tyr Glu Gln
85 90

<210> 32
<211> 200
<212> PRT
<213> SARS coronavirus

<400> 32
Met Phe Ile Phe Leu Leu Phe Leu Thr Leu Thr Ser Gly Ser Asp Leu
1 5 10 15
Asp Arg Cys Thr Thr Phe Asp Asp Val Gln Ala Pro Asn Tyr Thr Gln
20 25 30
His Thr Ser Ser Met Arg Gly Val Tyr Tyr Pro Asp Glu Ile Phe Arg
35 40 45
Ser Asp Thr Leu Tyr Leu Thr Gln Asp Leu Phe Leu Pro Phe Tyr Ser
50 55 60
Asn Val Thr Gly Phe His Thr Ile Asn His Thr Phe Gly Asn Pro Val
65 70 75 80
Ile Pro Phe Lys Asp Gly Ile Tyr Phe Ala Ala Thr Glu Lys Ser Asn
85 90 95
Val Val Arg Gly Trp Val Phe Gly Ser Thr Met Asn Asn Lys Ser Gln
100 105 110

Ser Val Ile Ile Ile Asn Asn Ser Thr Asn Val Val Ile Arg Ala Cys
115 120 125
Asn Phe Glu Leu Cys Asp Asn Pro Phe Phe Ala Val Ser Lys Pro Met
130 135 140
Gly Thr Gln Thr His Thr Met Ile Phe Asp Asn Ala Phe Asn Cys Thr
145 150 155 160
Phe Glu Tyr Ile Ser Asp Ala Phe Ser Leu Asp Val Ser Glu Lys Ser
165 170 175
Gly Asn Phe Lys His Leu Arg Glu Phe Val Phe Lys Asn Lys Asp Gly
180 185 190
Phe Leu Tyr Val Tyr Lys Gly Tyr
195 200

<210> 33
<211> 200
<212> PRT
<213> SARS coronavirus

<400> 33
Gln Pro Ile Asp Val Val Arg Asp Leu Pro Ser Gly Phe Asn Thr Leu
1 5 10 15
Lys Pro Ile Phe Lys Leu Pro Leu Gly Ile Asn Ile Thr Asn Phe Arg
20 25 30
Ala Ile Leu Thr Ala Phe Ser Pro Ala Gln Asp Ile Trp Gly Thr Ser
35 40 45
Ala Ala Ala Tyr Phe Val Gly Tyr Leu Lys Pro Thr Thr Phe Met Leu
50 55 60
Lys Tyr Asp Glu Asn Gly Thr Ile Thr Asp Ala Val Asp Cys Ser Gln
65 70 75 80
Asn Pro Leu Ala Glu Leu Lys Cys Ser Val Lys Ser Phe Glu Ile Asp
85 90 95
Lys Gly Ile Tyr Gln Thr Ser Asn Phe Arg Val Val Pro Ser Gly Asp
100 105 110
Val Val Arg Phe Pro Asn Ile Thr Asn Leu Cys Pro Phe Gly Glu Val
115 120 125
Phe Asn Ala Thr Lys Phe Pro Ser Val Tyr Ala Trp Glu Arg Lys Lys
130 135 140
Ile Ser Asn Cys Val Ala Asp Tyr Ser Val Leu Tyr Asn Ser Thr Phe
145 150 155 160
Phe Ser Thr Phe Lys Cys Tyr Gly Val Ser Ala Thr Lys Leu Asn Asp
165 170 175

Leu Cys Phe Ser Asn Val Tyr Ala Asp Ser Phe Val Val Lys Gly Asp
180 185 190

Asp Val Arg Gln Ile Ala Pro Gly
195 200

<210> 34
<211> 200
<212> PRT
<213> SARS coronavirus

<400> 34
Gln Thr Gly Val Ile Ala Asp Tyr Asn Tyr Lys Leu Pro Asp Asp Phe
1 5 10 15
Met Gly Cys Val Leu Ala Trp Asn Thr Arg Asn Ile Asp Ala Thr Ser
20 25 30
Thr Gly Asn Tyr Asn Tyr Lys Tyr Arg Tyr Leu Arg His Gly Lys Leu
35 40 45
Arg Pro Phe Glu Arg Asp Ile Ser Asn Val Pro Phe Ser Pro Asp Gly
50 55 60
Lys Pro Cys Thr Pro Pro Ala Leu Asn Cys Tyr Trp Pro Leu Asn Asp
65 70 75 80
Tyr Gly Phe Tyr Thr Thr Thr Gly Ile Gly Tyr Gln Pro Tyr Arg Val
85 90 95
Val Val Leu Ser Phe Glu Leu Leu Asn Ala Pro Ala Thr Val Cys Gly
100 105 110
Pro Lys Leu Ser Thr Asp Leu Ile Lys Asn Gln Cys Val Asn Phe Asn
115 120 125
Phe Asn Gly Leu Thr Gly Thr Gly Val Leu Thr Pro Ser Ser Lys Arg
130 135 140
Phe Gln Pro Phe Gln Gln Phe Gly Arg Asp Val Ser Asp Phe Thr Asp
145 150 155 160
Ser Val Arg Asp Pro Lys Thr Ser Glu Ile Leu Asp Ile Ser Pro Cys
165 170 175
Ala Phe Gly Gly Val Ser Val Ile Thr Pro Gly Thr Asn Ala Ser Ser
180 185 190
Glu Val Ala Val Leu Tyr Gln Asp
195 200

<210> 35
<211> 200
<212> PRT
<213> SARS coronavirus

<400> 35

Val Asn Cys Thr Asp Val Ser Thr Ala Ile His Ala Asp Gln Leu Thr
1 5 10 15
Pro Ala Trp Arg Ile Tyr Ser Thr Gly Asn Asn Val Phe Gln Thr Gln
20 25 30
Ala Gly Cys Leu Ile Gly Ala Glu His Val Asp Thr Ser Tyr Glu Cys
35 40 45
Asp Ile Pro Ile Gly Ala Gly Ile Cys Ala Ser Tyr His Thr Val Ser
50 55 60
Leu Leu Arg Ser Thr Ser Gln Lys Ser Ile Val Ala Tyr Thr Met Ser
65 70 75 80
Leu Gly Ala Asp Ser Ser Ile Ala Tyr Ser Asn Asn Thr Ile Ala Ile
85 90 95
Pro Thr Asn Phe Ser Ile Ser Ile Thr Thr Glu Val Met Pro Val Ser
100 105 110
Met Ala Lys Thr Ser Val Asp Cys Asn Met Tyr Ile Cys Gly Asp Ser
115 120 125
Thr Glu Cys Ala Asn Leu Leu Leu Gln Tyr Gly Ser Phe Cys Thr Gln
130 135 140
Leu Asn Arg Ala Leu Ser Gly Ile Ala Ala Glu Gln Asp Arg Asn Thr
145 150 155 160
Arg Glu Val Phe Ala Gln Val Lys Gln Met Tyr Lys Thr Pro Thr Leu
165 170 175
Lys Tyr Phe Gly Gly Phe Asn Phe Ser Gln Ile Leu Pro Asp Pro Leu
180 185 190
Lys Pro Thr Lys Arg Ser Phe Ile
195 200

<210> 36
<211> 200
<212> PRT
<213> SARS coronavirus

<400> 36

Glu Asp Leu Leu Phe Asn Lys Val Thr Leu Ala Asp Ala Gly Phe Met
1 5 10 15
Lys Gln Tyr Gly Glu Cys Leu Gly Asp Ile Asn Ala Arg Asp Leu Ile
20 25 30
Cys Ala Gln Lys Phe Asn Gly Leu Thr Val Leu Pro Pro Leu Leu Thr
35 40 45
Asp Asp Met Ile Ala Ala Tyr Thr Ala Ala Leu Val Ser Gly Thr Ala
50 55 60

Thr Ala Gly Trp Thr Phe Gly Ala Gly Ala Ala Leu Gln Ile Pro Phe
65 70 75 80
Ala Met Gln Met Ala Tyr Arg Phe Asn Gly Ile Gly Val Thr Gln Asn
85 90 95
Val Leu Tyr Glu Asn Gln Lys Gln Ile Ala Asn Gln Phe Asn Lys Ala
100 105 110
Ile Ser Gln Ile Gln Glu Ser Leu Thr Thr Thr Ser Thr Ala Leu Gly
115 120 125
Lys Leu Gln Asp Val Val Asn Gln Asn Ala Gln Ala Leu Asn Thr Leu
130 135 140
Val Lys Gln Leu Ser Ser Asn Phe Gly Ala Ile Ser Ser Val Leu Asn
145 150 155 160
Asp Ile Leu Ser Arg Leu Asp Lys Val Glu Ala Glu Val Gln Ile Asp
165 170 175
Arg Leu Ile Thr Gly Arg Leu Gln Ser Leu Gln Thr Tyr Val Thr Gln
180 185 190
Gln Leu Ile Arg Ala Ala Glu Ile
195 200

<210> 37
<211> 190
<212> PRT
<213> SARS coronavirus

<400> 37
Arg Ala Ser Ala Asn Leu Ala Ala Thr Lys Met Ser Glu Cys Val Leu
1 5 10 15
Gly Gln Ser Lys Arg Val Asp Phe Cys Gly Lys Gly Tyr His Leu Met
20 25 30
Ser Phe Pro Gln Ala Ala Pro His Gly Val Val Phe Leu His Val Thr
35 40 45
Tyr Val Pro Ser Gln Glu Arg Asn Phe Thr Thr Ala Pro Ala Ile Cys
50 55 60
His Glu Gly Lys Ala Tyr Phe Pro Arg Glu Gly Val Phe Val Phe Asn
65 70 75 80
Gly Thr Ser Trp Phe Ile Thr Gln Arg Asn Phe Phe Ser Pro Gln Ile
85 90 95
Ile Thr Thr Asp Asn Thr Phe Val Ser Gly Asn Cys Asp Val Val Ile
100 105 110
Gly Ile Ile Asn Asn Thr Val Tyr Asp Pro Leu Gln Pro Glu Leu Asp
115 120 125

Ser Phe Lys Glu Glu Leu Asp Lys Tyr Phe Lys Asn His Thr Ser Pro
130 135 140
Asp Val Asp Leu Gly Asp Ile Ser Gly Ile Asn Ala Ser Val Val Asn
145 150 155 160
Ile Gln Lys Glu Ile Asp Arg Leu Asn Glu Val Ala Lys Asn Leu Asn
165 170 175
Glu Ser Leu Ile Asp Leu Gln Glu Leu Gly Lys Tyr Glu Gln
180 185 190

<210> 38
<211> 400
<212> PRT
<213> SARS coronavirus

<400> 38
Met Phe Ile Phe Leu Leu Phe Leu Thr Leu Thr Ser Gly Ser Asp Leu
1 5 10 15
Asp Arg Cys Thr Thr Phe Asp Asp Val Gln Ala Pro Asn Tyr Thr Gln
20 25 30
His Thr Ser Ser Met Arg Gly Val Tyr Tyr Pro Asp Glu Ile Phe Arg
35 40 45
Ser Asp Thr Leu Tyr Leu Thr Gln Asp Leu Phe Leu Pro Phe Tyr Ser
50 55 60
Asn Val Thr Gly Phe His Thr Ile Asn His Thr Phe Gly Asn Pro Val
65 70 75 80
Ile Pro Phe Lys Asp Gly Ile Tyr Phe Ala Ala Thr Glu Lys Ser Asn
85 90 95
Val Val Arg Gly Trp Val Phe Gly Ser Thr Met Asn Asn Lys Ser Gln
100 105 110
Ser Val Ile Ile Ile Asn Asn Ser Thr Asn Val Val Ile Arg Ala Cys
115 120 125
Asn Phe Glu Leu Cys Asp Asn Pro Phe Phe Ala Val Ser Lys Pro Met
130 135 140
Gly Thr Gln Thr His Thr Met Ile Phe Asp Asn Ala Phe Asn Cys Thr
145 150 155 160
Phe Glu Tyr Ile Ser Asp Ala Phe Ser Leu Asp Val Ser Glu Lys Ser
165 170 175
Gly Asn Phe Lys His Leu Arg Glu Phe Val Phe Lys Asn Lys Asp Gly
180 185 190
Phe Leu Tyr Val Tyr Lys Gly Tyr Gln Pro Ile Asp Val Val Arg Asp
195 200 205

Leu Pro Ser Gly Phe Asn Thr Leu Lys Pro Ile Phe Lys Leu Pro Leu
210 215 220

Gly Ile Asn Ile Thr Asn Phe Arg Ala Ile Leu Thr Ala Phe Ser Pro
225 230 235 240

Ala Gln Asp Ile Trp Gly Thr Ser Ala Ala Ala Tyr Phe Val Gly Tyr
245 250 255

Leu Lys Pro Thr Thr Phe Met Leu Lys Tyr Asp Glu Asn Gly Thr Ile
260 265 270

Thr Asp Ala Val Asp Cys Ser Gln Asn Pro Leu Ala Glu Leu Lys Cys
275 280 285

Ser Val Lys Ser Phe Glu Ile Asp Lys Gly Ile Tyr Gln Thr Ser Asn
290 295 300

Phe Arg Val Val Pro Ser Gly Asp Val Val Arg Phe Pro Asn Ile Thr
305 310 315 320

Asn Leu Cys Pro Phe Gly Glu Val Phe Asn Ala Thr Lys Phe Pro Ser
325 330 335

Val Tyr Ala Trp Glu Arg Lys Lys Ile Ser Asn Cys Val Ala Asp Tyr
340 345 350

Ser Val Leu Tyr Asn Ser Thr Phe Phe Ser Thr Phe Lys Cys Tyr Gly
355 360 365

Val Ser Ala Thr Lys Leu Asn Asp Leu Cys Phe Ser Asn Val Tyr Ala
370 375 380

Asp Ser Phe Val Val Lys Gly Asp Asp Val Arg Gln Ile Ala Pro Gly
385 390 395 400

<210> 39

<211> 600

<212> PRT

<213> SARS coronavirus

<400> 39

Met Phe Ile Phe Leu Leu Phe Leu Thr Leu Thr Ser Gly Ser Asp Leu
1 5 10 15

Asp Arg Cys Thr Thr Phe Asp Asp Val Gln Ala Pro Asn Tyr Thr Gln
20 25 30

His Thr Ser Ser Met Arg Gly Val Tyr Tyr Pro Asp Glu Ile Phe Arg
35 40 45

Ser Asp Thr Leu Tyr Leu Thr Gln Asp Leu Phe Leu Pro Phe Tyr Ser
50 55 60

Asn Val Thr Gly Phe His Thr Ile Asn His Thr Phe Gly Asn Pro Val
65 70 75 80

Ile Pro Phe Lys Asp Gly Ile Tyr Phe Ala Ala Thr Glu Lys Ser Asn
85 90 95

Val Val Arg Gly Trp Val Phe Gly Ser Thr Met Asn Asn Lys Ser Gln
100 105 110

Ser Val Ile Ile Ile Asn Asn Ser Thr Asn Val Val Ile Arg Ala Cys
115 120 125

Asn Phe Glu Leu Cys Asp Asn Pro Phe Phe Ala Val Ser Lys Pro Met
130 135 140

Gly Thr Gln Thr His Thr Met Ile Phe Asp Asn Ala Phe Asn Cys Thr
145 150 155 160

Phe Glu Tyr Ile Ser Asp Ala Phe Ser Leu Asp Val Ser Glu Lys Ser
165 170 175

Gly Asn Phe Lys His Leu Arg Glu Phe Val Phe Lys Asn Lys Asp Gly
180 185 190

Phe Leu Tyr Val Tyr Lys Gly Tyr Gln Pro Ile Asp Val Val Arg Asp
195 200 205

Leu Pro Ser Gly Phe Asn Thr Leu Lys Pro Ile Phe Lys Leu Pro Leu
210 215 220

Gly Ile Asn Ile Thr Asn Phe Arg Ala Ile Leu Thr Ala Phe Ser Pro
225 230 235 240

Ala Gln Asp Ile Trp Gly Thr Ser Ala Ala Ala Tyr Phe Val Gly Tyr
245 250 255

Leu Lys Pro Thr Thr Phe Met Leu Lys Tyr Asp Glu Asn Gly Thr Ile
260 265 270

Thr Asp Ala Val Asp Cys Ser Gln Asn Pro Leu Ala Glu Leu Lys Cys
275 280 285

Ser Val Lys Ser Phe Glu Ile Asp Lys Gly Ile Tyr Gln Thr Ser Asn
290 295 300

Phe Arg Val Val Pro Ser Gly Asp Val Val Arg Phe Pro Asn Ile Thr
305 310 315 320

Asn Leu Cys Pro Phe Gly Glu Val Phe Asn Ala Thr Lys Phe Pro Ser
325 330 335

Val Tyr Ala Trp Glu Arg Lys Lys Ile Ser Asn Cys Val Ala Asp Tyr
340 345 350

Ser Val Leu Tyr Asn Ser Thr Phe Phe Ser Thr Phe Lys Cys Tyr Gly
355 360 365

Val Ser Ala Thr Lys Leu Asn Asp Leu Cys Phe Ser Asn Val Tyr Ala
370 375 380

Asp Ser Phe Val Val Lys Gly Asp Asp Val Arg Gln Ile Ala Pro Gly
385 390 395 400

Gln Thr Gly Val Ile Ala Asp Tyr Asn Tyr Lys Leu Pro Asp Asp Phe
405 410 415

Met Gly Cys Val Leu Ala Trp Asn Thr Arg Asn Ile Asp Ala Thr Ser
420 425 430

Thr Gly Asn Tyr Asn Tyr Lys Tyr Arg Tyr Leu Arg His Gly Lys Leu
435 440 445

Arg Pro Phe Glu Arg Asp Ile Ser Asn Val Pro Phe Ser Pro Asp Gly
450 455 460

Lys Pro Cys Thr Pro Pro Ala Leu Asn Cys Tyr Trp Pro Leu Asn Asp
465 470 475 480

Tyr Gly Phe Tyr Thr Thr Thr Gly Ile Gly Tyr Gln Pro Tyr Arg Val
485 490 495

Val Val Leu Ser Phe Glu Leu Leu Asn Ala Pro Ala Thr Val Cys Gly
500 505 510

Pro Lys Leu Ser Thr Asp Leu Ile Lys Asn Gln Cys Val Asn Phe Asn
515 520 525

Phe Asn Gly Leu Thr Gly Thr Gly Val Leu Thr Pro Ser Ser Lys Arg
530 535 540

Phe Gln Pro Phe Gln Gln Phe Gly Arg Asp Val Ser Asp Phe Thr Asp
545 550 555 560

Ser Val Arg Asp Pro Lys Thr Ser Glu Ile Leu Asp Ile Ser Pro Cys
565 570 575

Ala Phe Gly Gly Val Ser Val Ile Thr Pro Gly Thr Asn Ala Ser Ser
580 585 590

Glu Val Ala Val Leu Tyr Gln Asp
595 600

<210> 40
<211> 800
<212> PRT
<213> SARS coronavirus

<400> 40

Met Phe Ile Phe Leu Leu Phe Leu Thr Leu Thr Ser Gly Ser Asp Leu
1 5 10 15

Asp Arg Cys Thr Thr Phe Asp Asp Val Gln Ala Pro Asn Tyr Thr Gln
20 25 30

His Thr Ser Ser Met Arg Gly Val Tyr Tyr Pro Asp Glu Ile Phe Arg
35 40 45

Ser Asp Thr Leu Tyr Leu Thr Gln Asp Leu Phe Leu Pro Phe Tyr Ser
50 55 60

Asn Val Thr Gly Phe His Thr Ile Asn His Thr Phe Gly Asn Pro Val
65 70 75 80

Ile Pro Phe Lys Asp Gly Ile Tyr Phe Ala Ala Thr Glu Lys Ser Asn
85 90 95

Val Val Arg Gly Trp Val Phe Gly Ser Thr Met Asn Asn Lys Ser Gln
100 105 110

Ser Val Ile Ile Ile Asn Asn Ser Thr Asn Val Val Ile Arg Ala Cys
115 120 125

Asn Phe Glu Leu Cys Asp Asn Pro Phe Phe Ala Val Ser Lys Pro Met
130 135 140

Gly Thr Gln Thr His Thr Met Ile Phe Asp Asn Ala Phe Asn Cys Thr
145 150 155 160

Phe Glu Tyr Ile Ser Asp Ala Phe Ser Leu Asp Val Ser Glu Lys Ser
165 170 175

Gly Asn Phe Lys His Leu Arg Glu Phe Val Phe Lys Asn Lys Asp Gly
180 185 190

Phe Leu Tyr Val Tyr Lys Gly Tyr Gln Pro Ile Asp Val Val Arg Asp
195 200 205

Leu Pro Ser Gly Phe Asn Thr Leu Lys Pro Ile Phe Lys Leu Pro Leu
210 215 220

Gly Ile Asn Ile Thr Asn Phe Arg Ala Ile Leu Thr Ala Phe Ser Pro
225 230 235 240

Ala Gln Asp Ile Trp Gly Thr Ser Ala Ala Ala Tyr Phe Val Gly Tyr
245 250 255

Leu Lys Pro Thr Thr Phe Met Leu Lys Tyr Asp Glu Asn Gly Thr Ile
260 265 270

Thr Asp Ala Val Asp Cys Ser Gln Asn Pro Leu Ala Glu Leu Lys Cys
275 280 285

Ser Val Lys Ser Phe Glu Ile Asp Lys Gly Ile Tyr Gln Thr Ser Asn
290 295 300

Phe Arg Val Val Pro Ser Gly Asp Val Val Arg Phe Pro Asn Ile Thr
305 310 315 320

Asn Leu Cys Pro Phe Gly Glu Val Phe Asn Ala Thr Lys Phe Pro Ser
325 330 335

Val Tyr Ala Trp Glu Arg Lys Lys Ile Ser Asn Cys Val Ala Asp Tyr
340 345 350

Ser Val Leu Tyr Asn Ser Thr Phe Phe Ser Thr Phe Lys Cys Tyr Gly
355 360 365

Val Ser Ala Thr Lys Leu Asn Asp Leu Cys Phe Ser Asn Val Tyr Ala
370 375 380

Asp Ser Phe Val Val Lys Gly Asp Asp Val Arg Gln Ile Ala Pro Gly
385 390 395 400

Gln Thr Gly Val Ile Ala Asp Tyr Asn Tyr Lys Leu Pro Asp Asp Phe
405 410 415

Met Gly Cys Val Leu Ala Trp Asn Thr Arg Asn Ile Asp Ala Thr Ser
420 425 430

Thr Gly Asn Tyr Asn Tyr Lys Tyr Arg Tyr Leu Arg His Gly Lys Leu
435 440 445

Arg Pro Phe Glu Arg Asp Ile Ser Asn Val Pro Phe Ser Pro Asp Gly
450 455 460

Lys Pro Cys Thr Pro Pro Ala Leu Asn Cys Tyr Trp Pro Leu Asn Asp
465 470 475 480

Tyr Gly Phe Tyr Thr Thr Thr Gly Ile Gly Tyr Gln Pro Tyr Arg Val
485 490 495

Val Val Leu Ser Phe Glu Leu Leu Asn Ala Pro Ala Thr Val Cys Gly
500 505 510

Pro Lys Leu Ser Thr Asp Leu Ile Lys Asn Gln Cys Val Asn Phe Asn
515 520 525

Phe Asn Gly Leu Thr Gly Thr Gly Val Leu Thr Pro Ser Ser Lys Arg
530 535 540

Phe Gln Pro Phe Gln Gln Phe Gly Arg Asp Val Ser Asp Phe Thr Asp
545 550 555 560

Ser Val Arg Asp Pro Lys Thr Ser Glu Ile Leu Asp Ile Ser Pro Cys
565 570 575

Ala Phe Gly Gly Val Ser Val Ile Thr Pro Gly Thr Asn Ala Ser Ser
580 585 590

Glu Val Ala Val Leu Tyr Gln Asp Val Asn Cys Thr Asp Val Ser Thr
595 600 605

Ala Ile His Ala Asp Gln Leu Thr Pro Ala Trp Arg Ile Tyr Ser Thr
610 615 620

Gly Asn Asn Val Phe Gln Thr Gln Ala Gly Cys Leu Ile Gly Ala Glu
625 630 635 640

His Val Asp Thr Ser Tyr Glu Cys Asp Ile Pro Ile Gly Ala Gly Ile
645 650 655

Cys Ala Ser Tyr His Thr Val Ser Leu Leu Arg Ser Thr Ser Gln Lys
660 665 670

Ser Ile Val Ala Tyr Thr Met Ser Leu Gly Ala Asp Ser Ser Ile Ala
675 680 685

Tyr Ser Asn Asn Thr Ile Ala Ile Pro Thr Asn Phe Ser Ile Ser Ile
690 695 700

Thr Thr Glu Val Met Pro Val Ser Met Ala Lys Thr Ser Val Asp Cys
705 710 715 720

Asn Met Tyr Ile Cys Gly Asp Ser Thr Glu Cys Ala Asn Leu Leu Leu
725 730 735

Gln Tyr Gly Ser Phe Cys Thr Gln Leu Asn Arg Ala Leu Ser Gly Ile
740 745 750

Ala Ala Glu Gln Asp Arg Asn Thr Arg Glu Val Phe Ala Gln Val Lys
755 760 765
Gln Met Tyr Lys Thr Pro Thr Leu Lys Tyr Phe Gly Gly Phe Asn Phe
770 775 780
Ser Gln Ile Leu Pro Asp Pro Leu Lys Pro Thr Lys Arg Ser Phe Ile
785 790 795 800

<210> 41
<211> 1000
<212> PRT
<213> SARS coronavirus

<400> 41
Met Phe Ile Phe Leu Leu Phe Leu Thr Leu Thr Ser Gly Ser Asp Leu
1 5 10 15
Asp Arg Cys Thr Thr Phe Asp Asp Val Gln Ala Pro Asn Tyr Thr Gln
20 25 30
His Thr Ser Ser Met Arg Gly Val Tyr Tyr Pro Asp Glu Ile Phe Arg
35 40 45
Ser Asp Thr Leu Tyr Leu Thr Gln Asp Leu Phe Leu Pro Phe Tyr Ser
50 55 60
Asn Val Thr Gly Phe His Thr Ile Asn His Thr Phe Gly Asn Pro Val
65 70 75 80
Ile Pro Phe Lys Asp Gly Ile Tyr Phe Ala Ala Thr Glu Lys Ser Asn
85 90 95
Val Val Arg Gly Trp Val Phe Gly Ser Thr Met Asn Asn Lys Ser Gln
100 105 110
Ser Val Ile Ile Ile Asn Asn Ser Thr Asn Val Val Ile Arg Ala Cys
115 120 125
Asn Phe Glu Leu Cys Asp Asn Pro Phe Phe Ala Val Ser Lys Pro Met
130 135 140
Gly Thr Gln Thr His Thr Met Ile Phe Asp Asn Ala Phe Asn Cys Thr
145 150 155 160
Phe Glu Tyr Ile Ser Asp Ala Phe Ser Leu Asp Val Ser Glu Lys Ser
165 170 175
Gly Asn Phe Lys His Leu Arg Glu Phe Val Phe Lys Asn Lys Asp Gly
180 185 190
Phe Leu Tyr Val Tyr Lys Gly Tyr Gln Pro Ile Asp Val Val Arg Asp
195 200 205
Leu Pro Ser Gly Phe Asn Thr Leu Lys Pro Ile Phe Lys Leu Pro Leu
210 215 220

Gly Ile Asn Ile Thr Asn Phe Arg Ala Ile Leu Thr Ala Phe Ser Pro
225 230 235 240

Ala Gln Asp Ile Trp Gly Thr Ser Ala Ala Ala Tyr Phe Val Gly Tyr
245 250 255

Leu Lys Pro Thr Thr Phe Met Leu Lys Tyr Asp Glu Asn Gly Thr Ile
260 265 270

Thr Asp Ala Val Asp Cys Ser Gln Asn Pro Leu Ala Glu Leu Lys Cys
275 280 285

Ser Val Lys Ser Phe Glu Ile Asp Lys Gly Ile Tyr Gln Thr Ser Asn
290 295 300

Phe Arg Val Val Pro Ser Gly Asp Val Val Arg Phe Pro Asn Ile Thr
305 310 315 320

Asn Leu Cys Pro Phe Gly Glu Val Phe Asn Ala Thr Lys Phe Pro Ser
325 330 335

Val Tyr Ala Trp Glu Arg Lys Lys Ile Ser Asn Cys Val Ala Asp Tyr
340 345 350

Ser Val Leu Tyr Asn Ser Thr Phe Phe Ser Thr Phe Lys Cys Tyr Gly
355 360 365

Val Ser Ala Thr Lys Leu Asn Asp Leu Cys Phe Ser Asn Val Tyr Ala
370 375 380

Asp Ser Phe Val Val Lys Gly Asp Asp Val Arg Gln Ile Ala Pro Gly
385 390 395 400

Gln Thr Gly Val Ile Ala Asp Tyr Asn Tyr Lys Leu Pro Asp Asp Phe
405 410 415

Met Gly Cys Val Leu Ala Trp Asn Thr Arg Asn Ile Asp Ala Thr Ser
420 425 430

Thr Gly Asn Tyr Asn Tyr Lys Tyr Arg Tyr Leu Arg His Gly Lys Leu
435 440 445

Arg Pro Phe Glu Arg Asp Ile Ser Asn Val Pro Phe Ser Pro Asp Gly
450 455 460

Lys Pro Cys Thr Pro Pro Ala Leu Asn Cys Tyr Trp Pro Leu Asn Asp
465 470 475 480

Tyr Gly Phe Tyr Thr Thr Thr Gly Ile Gly Tyr Gln Pro Tyr Arg Val
485 490 495

Val Val Leu Ser Phe Glu Leu Leu Asn Ala Pro Ala Thr Val Cys Gly
500 505 510

Pro Lys Leu Ser Thr Asp Leu Ile Lys Asn Gln Cys Val Asn Phe Asn
515 520 525

Phe Asn Gly Leu Thr Gly Thr Gly Val Leu Thr Pro Ser Ser Lys Arg
530 535 540

Phe Gln Pro Phe Gln Gln Phe Gly Arg Asp Val Ser Asp Phe Thr Asp
545 550 555 560

Ser Val Arg Asp Pro Lys Thr Ser Glu Ile Leu Asp Ile Ser Pro Cys
565 570 575

Ala Phe Gly Gly Val Ser Val Ile Thr Pro Gly Thr Asn Ala Ser Ser
580 585 590

Glu Val Ala Val Leu Tyr Gln Asp Val Asn Cys Thr Asp Val Ser Thr
595 600 605

Ala Ile His Ala Asp Gln Leu Thr Pro Ala Trp Arg Ile Tyr Ser Thr
610 615 620

Gly Asn Asn Val Phe Gln Thr Gln Ala Gly Cys Leu Ile Gly Ala Glu
625 630 635 640

His Val Asp Thr Ser Tyr Glu Cys Asp Ile Pro Ile Gly Ala Gly Ile
645 650 655

Cys Ala Ser Tyr His Thr Val Ser Leu Leu Arg Ser Thr Ser Gln Lys
660 665 670

Ser Ile Val Ala Tyr Thr Met Ser Leu Gly Ala Asp Ser Ser Ile Ala
675 680 685

Tyr Ser Asn Asn Thr Ile Ala Ile Pro Thr Asn Phe Ser Ile Ser Ile
690 695 700

Thr Thr Glu Val Met Pro Val Ser Met Ala Lys Thr Ser Val Asp Cys
705 710 715 720

Asn Met Tyr Ile Cys Gly Asp Ser Thr Glu Cys Ala Asn Leu Leu Leu
725 730 735

Gln Tyr Gly Ser Phe Cys Thr Gln Leu Asn Arg Ala Leu Ser Gly Ile
740 745 750

Ala Ala Glu Gln Asp Arg Asn Thr Arg Glu Val Phe Ala Gln Val Lys
755 760 765

Gln Met Tyr Lys Thr Pro Thr Leu Lys Tyr Phe Gly Gly Phe Asn Phe
770 775 780

Ser Gln Ile Leu Pro Asp Pro Leu Lys Pro Thr Lys Arg Ser Phe Ile
785 790 795 800

Glu Asp Leu Leu Phe Asn Lys Val Thr Leu Ala Asp Ala Gly Phe Met
805 810 815

Lys Gln Tyr Gly Glu Cys Leu Gly Asp Ile Asn Ala Arg Asp Leu Ile
820 825 830

Cys Ala Gln Lys Phe Asn Gly Leu Thr Val Leu Pro Pro Leu Leu Thr
835 840 845

Asp Asp Met Ile Ala Ala Tyr Thr Ala Ala Leu Val Ser Gly Thr Ala
850 855 860

Thr Ala Gly Trp Thr Phe Gly Ala Gly Ala Ala Leu Gln Ile Pro Phe
865 870 875 880

Ala Met Gln Met Ala Tyr Arg Phe Asn Gly Ile Gly Val Thr Gln Asn
885 890 895

Val Leu Tyr Glu Asn Gln Lys Gln Ile Ala Asn Gln Phe Asn Lys Ala
900 905 910

Ile Ser Gln Ile Gln Glu Ser Leu Thr Thr Thr Ser Thr Ala Leu Gly
915 920 925

Lys Leu Gln Asp Val Val Asn Gln Asn Ala Gln Ala Leu Asn Thr Leu
930 935 940

Val Lys Gln Leu Ser Ser Asn Phe Gly Ala Ile Ser Ser Val Leu Asn
945 950 955 960

Asp Ile Leu Ser Arg Leu Asp Lys Val Glu Ala Glu Val Gln Ile Asp
965 970 975

Arg Leu Ile Thr Gly Arg Leu Gln Ser Leu Gln Thr Tyr Val Thr Gln
980 985 990

Gln Leu Ile Arg Ala Ala Glu Ile
995 1000

<210> 42

<211> 1190

<212> PRT

<213> SARS coronavirus

<400> 42

Met Phe Ile Phe Leu Leu Phe Leu Thr Leu Thr Ser Gly Ser Asp Leu
1 5 10 15

Asp Arg Cys Thr Thr Phe Asp Asp Val Gln Ala Pro Asn Tyr Thr Gln
20 25 30

His Thr Ser Ser Met Arg Gly Val Tyr Tyr Pro Asp Glu Ile Phe Arg
35 40 45

Ser Asp Thr Leu Tyr Leu Thr Gln Asp Leu Phe Leu Pro Phe Tyr Ser
50 55 60

Asn Val Thr Gly Phe His Thr Ile Asn His Thr Phe Gly Asn Pro Val
65 70 75 80

Ile Pro Phe Lys Asp Gly Ile Tyr Phe Ala Ala Thr Glu Lys Ser Asn
85 90 95

Val Val Arg Gly Trp Val Phe Gly Ser Thr Met Asn Asn Lys Ser Gln
100 105 110

Ser Val Ile Ile Ile Asn Asn Ser Thr Asn Val Val Ile Arg Ala Cys
115 120 125

Asn Phe Glu Leu Cys Asp Asn Pro Phe Phe Ala Val Ser Lys Pro Met
130 135 140

Gly Thr Gln Thr His Thr Met Ile Phe Asp Asn Ala Phe Asn Cys Thr
145 150 155 160

Phe Glu Tyr Ile Ser Asp Ala Phe Ser Leu Asp Val Ser Glu Lys Ser
165 170 175

Gly Asn Phe Lys His Leu Arg Glu Phe Val Phe Lys Asn Lys Asp Gly
180 185 190

Phe Leu Tyr Val Tyr Lys Gly Tyr Gln Pro Ile Asp Val Val Arg Asp
195 200 205

Leu Pro Ser Gly Phe Asn Thr Leu Lys Pro Ile Phe Lys Leu Pro Leu
210 215 220

Gly Ile Asn Ile Thr Asn Phe Arg Ala Ile Leu Thr Ala Phe Ser Pro
225 230 235 240

Ala Gln Asp Ile Trp Gly Thr Ser Ala Ala Ala Tyr Phe Val Gly Tyr
245 250 255

Leu Lys Pro Thr Thr Phe Met Leu Lys Tyr Asp Glu Asn Gly Thr Ile
260 265 270

Thr Asp Ala Val Asp Cys Ser Gln Asn Pro Leu Ala Glu Leu Lys Cys
275 280 285

Ser Val Lys Ser Phe Glu Ile Asp Lys Gly Ile Tyr Gln Thr Ser Asn
290 295 300

Phe Arg Val Val Pro Ser Gly Asp Val Val Arg Phe Pro Asn Ile Thr
305 310 315 320

Asn Leu Cys Pro Phe Gly Glu Val Phe Asn Ala Thr Lys Phe Pro Ser
325 330 335

Val Tyr Ala Trp Glu Arg Lys Lys Ile Ser Asn Cys Val Ala Asp Tyr
340 345 350

Ser Val Leu Tyr Asn Ser Thr Phe Phe Ser Thr Phe Lys Cys Tyr Gly
355 360 365

Val Ser Ala Thr Lys Leu Asn Asp Leu Cys Phe Ser Asn Val Tyr Ala
370 375 380

Asp Ser Phe Val Val Lys Gly Asp Asp Val Arg Gln Ile Ala Pro Gly
385 390 395 400

Gln Thr Gly Val Ile Ala Asp Tyr Asn Tyr Lys Leu Pro Asp Asp Phe
405 410 415

Met Gly Cys Val Leu Ala Trp Asn Thr Arg Asn Ile Asp Ala Thr Ser
420 425 430

Thr Gly Asn Tyr Asn Tyr Lys Tyr Arg Tyr Leu Arg His Gly Lys Leu
435 440 445

Arg Pro Phe Glu Arg Asp Ile Ser Asn Val Pro Phe Ser Pro Asp Gly
450 455 460

Lys Pro Cys Thr Pro Pro Ala Leu Asn Cys Tyr Trp Pro Leu Asn Asp
465 470 475 480

Tyr Gly Phe Tyr Thr Thr Thr Gly Ile Gly Tyr Gln Pro Tyr Arg Val
485 490 495

Val Val Leu Ser Phe Glu Leu Leu Asn Ala Pro Ala Thr Val Cys Gly
500 505 510

Pro Lys Leu Ser Thr Asp Leu Ile Lys Asn Gln Cys Val Asn Phe Asn
515 520 525

Phe Asn Gly Leu Thr Gly Thr Gly Val Leu Thr Pro Ser Ser Lys Arg
530 535 540

Phe Gln Pro Phe Gln Gln Phe Gly Arg Asp Val Ser Asp Phe Thr Asp
545 550 555 560

Ser Val Arg Asp Pro Lys Thr Ser Glu Ile Leu Asp Ile Ser Pro Cys
565 570 575

Ala Phe Gly Gly Val Ser Val Ile Thr Pro Gly Thr Asn Ala Ser Ser
580 585 590

Glu Val Ala Val Leu Tyr Gln Asp Val Asn Cys Thr Asp Val Ser Thr
595 600 605

Ala Ile His Ala Asp Gln Leu Thr Pro Ala Trp Arg Ile Tyr Ser Thr
610 615 620

Gly Asn Asn Val Phe Gln Thr Gln Ala Gly Cys Leu Ile Gly Ala Glu
625 630 635 640

His Val Asp Thr Ser Tyr Glu Cys Asp Ile Pro Ile Gly Ala Gly Ile
645 650 655

Cys Ala Ser Tyr His Thr Val Ser Leu Leu Arg Ser Thr Ser Gln Lys
660 665 670

Ser Ile Val Ala Tyr Thr Met Ser Leu Gly Ala Asp Ser Ser Ile Ala
675 680 685

Tyr Ser Asn Asn Thr Ile Ala Ile Pro Thr Asn Phe Ser Ile Ser Ile
690 695 700

Thr Thr Glu Val Met Pro Val Ser Met Ala Lys Thr Ser Val Asp Cys
705 710 715 720

Asn Met Tyr Ile Cys Gly Asp Ser Thr Glu Cys Ala Asn Leu Leu Leu
725 730 735

Gln Tyr Gly Ser Phe Cys Thr Gln Leu Asn Arg Ala Leu Ser Gly Ile
740 745 750

Ala Ala Glu Gln Asp Arg Asn Thr Arg Glu Val Phe Ala Gln Val Lys
755 760 765

Gln Met Tyr Lys Thr Pro Thr Leu Lys Tyr Phe Gly Gly Phe Asn Phe
770 775 780

Ser Gln Ile Leu Pro Asp Pro Leu Lys Pro Thr Lys Arg Ser Phe Ile
785 790 795 800

Glu Asp Leu Leu Phe Asn Lys Val Thr Leu Ala Asp Ala Gly Phe Met
805 810 815

Lys Gln Tyr Gly Glu Cys Leu Gly Asp Ile Asn Ala Arg Asp Leu Ile
820 825 830

Cys Ala Gln Lys Phe Asn Gly Leu Thr Val Leu Pro Pro Leu Leu Thr
835 840 845

Asp Asp Met Ile Ala Ala Tyr Thr Ala Ala Leu Val Ser Gly Thr Ala
850 855 860

Thr Ala Gly Trp Thr Phe Gly Ala Gly Ala Ala Leu Gln Ile Pro Phe
865 870 875 880

Ala Met Gln Met Ala Tyr Arg Phe Asn Gly Ile Gly Val Thr Gln Asn
885 890 895

Val Leu Tyr Glu Asn Gln Lys Gln Ile Ala Asn Gln Phe Asn Lys Ala
900 905 910

Ile Ser Gln Ile Gln Glu Ser Leu Thr Thr Thr Ser Thr Ala Leu Gly
915 920 925

Lys Leu Gln Asp Val Val Asn Gln Asn Ala Gln Ala Leu Asn Thr Leu
930 935 940

Val Lys Gln Leu Ser Ser Asn Phe Gly Ala Ile Ser Ser Val Leu Asn
945 950 955 960

Asp Ile Leu Ser Arg Leu Asp Lys Val Glu Ala Glu Val Gln Ile Asp
965 970 975

Arg Leu Ile Thr Gly Arg Leu Gln Ser Leu Gln Thr Tyr Val Thr Gln
980 985 990

Gln Leu Ile Arg Ala Ala Glu Ile Arg Ala Ser Ala Asn Leu Ala Ala
995 1000 1005

Thr Lys Met Ser Glu Cys Val Leu Gly Gln Ser Lys Arg Val Asp Phe
1010 1015 1020

Cys Gly Lys Gly Tyr His Leu Met Ser Phe Pro Gln Ala Ala Pro His
1025 1030 1035 1040

Gly Val Val Phe Leu His Val Thr Tyr Val Pro Ser Gln Glu Arg Asn
1045 1050 1055

Phe Thr Thr Ala Pro Ala Ile Cys His Glu Gly Lys Ala Tyr Phe Pro
1060 1065 1070

Arg Glu Gly Val Phe Val Phe Asn Gly Thr Ser Trp Phe Ile Thr Gln
1075 1080 1085

Arg Asn Phe Phe Ser Pro Gln Ile Ile Thr Thr Asp Asn Thr Phe Val
1090 1095 1100

Ser Gly Asn Cys Asp Val Val Ile Gly Ile Ile Asn Asn Thr Val Tyr
1105 1110 1115 1120

Asp Pro Leu Gln Pro Glu Leu Asp Ser Phe Lys Glu Glu Leu Asp Lys
1125 1130 1135

Tyr Phe Lys Asn His Thr Ser Pro Asp Val Asp Leu Gly Asp Ile Ser
1140 1145 1150

Gly Ile Asn Ala Ser Val Val Asn Ile Gln Lys Glu Ile Asp Arg Leu
1155 1160 1165

Asn Glu Val Ala Lys Asn Leu Asn Glu Ser Leu Ile Asp Leu Gln Glu
1170 1175 1180

Leu Gly Lys Tyr Glu Gln
1185 1190

<210> 43
<211> 84
<212> PRT
<213> SARS coronavirus

<400> 43
Asp Arg Cys Thr Thr Phe Asp Asp Val Gln Ala Pro Asn Tyr Thr Gln
1 5 10 15
His Thr Ser Ser Met Arg Gly Val Tyr Tyr Pro Asp Glu Ile Phe Arg
20 25 30
Ser Asp Thr Leu Tyr Leu Thr Gln Asp Leu Phe Leu Pro Phe Tyr Ser
35 40 45
Asn Val Thr Gly Phe His Thr Ile Asn His Thr Phe Gly Asn Pro Val
50 55 60
Ile Pro Phe Lys Asp Gly Ile Tyr Phe Ala Ala Thr Glu Lys Ser Asn
65 70 75 80
Val Val Arg Gly

<210> 44
<211> 184
<212> PRT
<213> SARS coronavirus

<400> 44
Asp Arg Cys Thr Thr Phe Asp Asp Val Gln Ala Pro Asn Tyr Thr Gln
1 5 10 15
His Thr Ser Ser Met Arg Gly Val Tyr Tyr Pro Asp Glu Ile Phe Arg
20 25 30
Ser Asp Thr Leu Tyr Leu Thr Gln Asp Leu Phe Leu Pro Phe Tyr Ser
35 40 45
Asn Val Thr Gly Phe His Thr Ile Asn His Thr Phe Gly Asn Pro Val
50 55 60
Ile Pro Phe Lys Asp Gly Ile Tyr Phe Ala Ala Thr Glu Lys Ser Asn
65 70 75 80
Val Val Arg Gly Trp Val Phe Gly Ser Thr Met Asn Asn Lys Ser Gln
85 90 95

```
Ser Val Ile Ile Ile Asn Asn Ser Thr Asn Val Val Ile Arg Ala Cys
            100                 105                 110
Asn Phe Glu Leu Cys Asp Asn Pro Phe Phe Ala Val Ser Lys Pro Met
        115                 120                 125
Gly Thr Gln Thr His Thr Met Ile Phe Asp Asn Ala Phe Asn Cys Thr
    130                 135                 140
Phe Glu Tyr Ile Ser Asp Ala Phe Ser Leu Asp Val Ser Glu Lys Ser
145                 150                 155                 160
Gly Asn Phe Lys His Leu Arg Glu Phe Val Phe Lys Asn Lys Asp Gly
                165                 170                 175
Phe Leu Tyr Val Tyr Lys Gly Tyr
            180
```

<210> 45
<211> 384
<212> PRT
<213> SARS coronavirus

<400> 45

```
Asp Arg Cys Thr Thr Phe Asp Asp Val Gln Ala Pro Asn Tyr Thr Gln
1               5                   10                  15
His Thr Ser Ser Met Arg Gly Val Tyr Tyr Pro Asp Glu Ile Phe Arg
            20                  25                  30
Ser Asp Thr Leu Tyr Leu Thr Gln Asp Leu Phe Leu Pro Phe Tyr Ser
        35                  40                  45
Asn Val Thr Gly Phe His Thr Ile Asn His Thr Phe Gly Asn Pro Val
    50                  55                  60
Ile Pro Phe Lys Asp Gly Ile Tyr Phe Ala Ala Thr Glu Lys Ser Asn
65                  70                  75                  80
Val Val Arg Gly Trp Val Phe Gly Ser Thr Met Asn Asn Lys Ser Gln
                85                  90                  95
Ser Val Ile Ile Ile Asn Asn Ser Thr Asn Val Val Ile Arg Ala Cys
            100                 105                 110
Asn Phe Glu Leu Cys Asp Asn Pro Phe Phe Ala Val Ser Lys Pro Met
        115                 120                 125
Gly Thr Gln Thr His Thr Met Ile Phe Asp Asn Ala Phe Asn Cys Thr
    130                 135                 140
Phe Glu Tyr Ile Ser Asp Ala Phe Ser Leu Asp Val Ser Glu Lys Ser
145                 150                 155                 160
Gly Asn Phe Lys His Leu Arg Glu Phe Val Phe Lys Asn Lys Asp Gly
                165                 170                 175
```

Phe Leu Tyr Val Tyr Lys Gly Tyr Gln Pro Ile Asp Val Val Arg Asp
180 185 190

Leu Pro Ser Gly Phe Asn Thr Leu Lys Pro Ile Phe Lys Leu Pro Leu
195 200 205

Gly Ile Asn Ile Thr Asn Phe Arg Ala Ile Leu Thr Ala Phe Ser Pro
210 215 220

Ala Gln Asp Ile Trp Gly Thr Ser Ala Ala Ala Tyr Phe Val Gly Tyr
225 230 235 240

Leu Lys Pro Thr Thr Phe Met Leu Lys Tyr Asp Glu Asn Gly Thr Ile
245 250 255

Thr Asp Ala Val Asp Cys Ser Gln Asn Pro Leu Ala Glu Leu Lys Cys
260 265 270

Ser Val Lys Ser Phe Glu Ile Asp Lys Gly Ile Tyr Gln Thr Ser Asn
275 280 285

Phe Arg Val Val Pro Ser Gly Asp Val Val Arg Phe Pro Asn Ile Thr
290 295 300

Asn Leu Cys Pro Phe Gly Glu Val Phe Asn Ala Thr Lys Phe Pro Ser
305 310 315 320

Val Tyr Ala Trp Glu Arg Lys Lys Ile Ser Asn Cys Val Ala Asp Tyr
325 330 335

Ser Val Leu Tyr Asn Ser Thr Phe Phe Ser Thr Phe Lys Cys Tyr Gly
340 345 350

Val Ser Ala Thr Lys Leu Asn Asp Leu Cys Phe Ser Asn Val Tyr Ala
355 360 365

Asp Ser Phe Val Val Lys Gly Asp Asp Val Arg Gln Ile Ala Pro Gly
370 375 380

<210> 46
<211> 584
<212> PRT
<213> SARS coronavirus

<400> 46

Asp Arg Cys Thr Thr Phe Asp Asp Val Gln Ala Pro Asn Tyr Thr Gln
1 5 10 15

His Thr Ser Ser Met Arg Gly Val Tyr Tyr Pro Asp Glu Ile Phe Arg
20 25 30

Ser Asp Thr Leu Tyr Leu Thr Gln Asp Leu Phe Leu Pro Phe Tyr Ser
35 40 45

Asn Val Thr Gly Phe His Thr Ile Asn His Thr Phe Gly Asn Pro Val
50 55 60

Ile Pro Phe Lys Asp Gly Ile Tyr Phe Ala Ala Thr Glu Lys Ser Asn
65 70 75 80

Val Val Arg Gly Trp Val Phe Gly Ser Thr Met Asn Asn Lys Ser Gln
85 90 95

Ser Val Ile Ile Ile Asn Asn Ser Thr Asn Val Val Ile Arg Ala Cys
100 105 110

Asn Phe Glu Leu Cys Asp Asn Pro Phe Phe Ala Val Ser Lys Pro Met
115 120 125

Gly Thr Gln Thr His Thr Met Ile Phe Asp Asn Ala Phe Asn Cys Thr
130 135 140

Phe Glu Tyr Ile Ser Asp Ala Phe Ser Leu Asp Val Ser Glu Lys Ser
145 150 155 160

Gly Asn Phe Lys His Leu Arg Glu Phe Val Phe Lys Asn Lys Asp Gly
165 170 175

Phe Leu Tyr Val Tyr Lys Gly Tyr Gln Pro Ile Asp Val Val Arg Asp
180 185 190

Leu Pro Ser Gly Phe Asn Thr Leu Lys Pro Ile Phe Lys Leu Pro Leu
195 200 205

Gly Ile Asn Ile Thr Asn Phe Arg Ala Ile Leu Thr Ala Phe Ser Pro
210 215 220

Ala Gln Asp Ile Trp Gly Thr Ser Ala Ala Ala Tyr Phe Val Gly Tyr
225 230 235 240

Leu Lys Pro Thr Thr Phe Met Leu Lys Tyr Asp Glu Asn Gly Thr Ile
245 250 255

Thr Asp Ala Val Asp Cys Ser Gln Asn Pro Leu Ala Glu Leu Lys Cys
260 265 270

Ser Val Lys Ser Phe Glu Ile Asp Lys Gly Ile Tyr Gln Thr Ser Asn
275 280 285

Phe Arg Val Val Pro Ser Gly Asp Val Val Arg Phe Pro Asn Ile Thr
290 295 300

Asn Leu Cys Pro Phe Gly Glu Val Phe Asn Ala Thr Lys Phe Pro Ser
305 310 315 320

Val Tyr Ala Trp Glu Arg Lys Lys Ile Ser Asn Cys Val Ala Asp Tyr
325 330 335

Ser Val Leu Tyr Asn Ser Thr Phe Phe Ser Thr Phe Lys Cys Tyr Gly
340 345 350

Val Ser Ala Thr Lys Leu Asn Asp Leu Cys Phe Ser Asn Val Tyr Ala
355 360 365

Asp Ser Phe Val Val Lys Gly Asp Asp Val Arg Gln Ile Ala Pro Gly
370 375 380

Gln Thr Gly Val Ile Ala Asp Tyr Asn Tyr Lys Leu Pro Asp Asp Phe
385 390 395 400

Met Gly Cys Val Leu Ala Trp Asn Thr Arg Asn Ile Asp Ala Thr Ser
405 410 415
Thr Gly Asn Tyr Asn Tyr Lys Tyr Arg Tyr Leu Arg His Gly Lys Leu
420 425 430
Arg Pro Phe Glu Arg Asp Ile Ser Asn Val Pro Phe Ser Pro Asp Gly
435 440 445
Lys Pro Cys Thr Pro Pro Ala Leu Asn Cys Tyr Trp Pro Leu Asn Asp
450 455 460
Tyr Gly Phe Tyr Thr Thr Thr Gly Ile Gly Tyr Gln Pro Tyr Arg Val
465 470 475 480
Val Val Leu Ser Phe Glu Leu Leu Asn Ala Pro Ala Thr Val Cys Gly
485 490 495
Pro Lys Leu Ser Thr Asp Leu Ile Lys Asn Gln Cys Val Asn Phe Asn
500 505 510
Phe Asn Gly Leu Thr Gly Thr Gly Val Leu Thr Pro Ser Ser Lys Arg
515 520 525
Phe Gln Pro Phe Gln Gln Phe Gly Arg Asp Val Ser Asp Phe Thr Asp
530 535 540
Ser Val Arg Asp Pro Lys Thr Ser Glu Ile Leu Asp Ile Ser Pro Cys
545 550 555 560
Ala Phe Gly Gly Val Ser Val Ile Thr Pro Gly Thr Asn Ala Ser Ser
565 570 575
Glu Val Ala Val Leu Tyr Gln Asp
580

<210> 47
<211> 784
<212> PRT
<213> SARS coronavirus

<400> 47
Asp Arg Cys Thr Thr Phe Asp Asp Val Gln Ala Pro Asn Tyr Thr Gln
1 5 10 15
His Thr Ser Ser Met Arg Gly Val Tyr Tyr Pro Asp Glu Ile Phe Arg
20 25 30
Ser Asp Thr Leu Tyr Leu Thr Gln Asp Leu Phe Leu Pro Phe Tyr Ser
35 40 45
Asn Val Thr Gly Phe His Thr Ile Asn His Thr Phe Gly Asn Pro Val
50 55 60
Ile Pro Phe Lys Asp Gly Ile Tyr Phe Ala Ala Thr Glu Lys Ser Asn
65 70 75 80

```
Val Val Arg Gly Trp Val Phe Gly Ser Thr Met Asn Asn Lys Ser Gln
                85                  90                  95
Ser Val Ile Ile Ile Asn Asn Ser Thr Asn Val Val Ile Arg Ala Cys
            100                 105                 110
Asn Phe Glu Leu Cys Asp Asn Pro Phe Phe Ala Val Ser Lys Pro Met
        115                 120                 125
Gly Thr Gln Thr His Thr Met Ile Phe Asp Asn Ala Phe Asn Cys Thr
    130                 135                 140
Phe Glu Tyr Ile Ser Asp Ala Phe Ser Leu Asp Val Ser Glu Lys Ser
145                 150                 155                 160
Gly Asn Phe Lys His Leu Arg Glu Phe Val Phe Lys Asn Lys Asp Gly
                165                 170                 175
Phe Leu Tyr Val Tyr Lys Gly Tyr Gln Pro Ile Asp Val Val Arg Asp
            180                 185                 190
Leu Pro Ser Gly Phe Asn Thr Leu Lys Pro Ile Phe Lys Leu Pro Leu
        195                 200                 205
Gly Ile Asn Ile Thr Asn Phe Arg Ala Ile Leu Thr Ala Phe Ser Pro
    210                 215                 220
Ala Gln Asp Ile Trp Gly Thr Ser Ala Ala Ala Tyr Phe Val Gly Tyr
225                 230                 235                 240
Leu Lys Pro Thr Thr Phe Met Leu Lys Tyr Asp Glu Asn Gly Thr Ile
                245                 250                 255
Thr Asp Ala Val Asp Cys Ser Gln Asn Pro Leu Ala Glu Leu Lys Cys
            260                 265                 270
Ser Val Lys Ser Phe Glu Ile Asp Lys Gly Ile Tyr Gln Thr Ser Asn
        275                 280                 285
Phe Arg Val Val Pro Ser Gly Asp Val Val Arg Phe Pro Asn Ile Thr
    290                 295                 300
Asn Leu Cys Pro Phe Gly Glu Val Phe Asn Ala Thr Lys Phe Pro Ser
305                 310                 315                 320
Val Tyr Ala Trp Glu Arg Lys Lys Ile Ser Asn Cys Val Ala Asp Tyr
                325                 330                 335
Ser Val Leu Tyr Asn Ser Thr Phe Phe Ser Thr Phe Lys Cys Tyr Gly
            340                 345                 350
Val Ser Ala Thr Lys Leu Asn Asp Leu Cys Phe Ser Asn Val Tyr Ala
        355                 360                 365
Asp Ser Phe Val Val Lys Gly Asp Asp Val Arg Gln Ile Ala Pro Gly
    370                 375                 380
Gln Thr Gly Val Ile Ala Asp Tyr Asn Tyr Lys Leu Pro Asp Asp Phe
385                 390                 395                 400
Met Gly Cys Val Leu Ala Trp Asn Thr Arg Asn Ile Asp Ala Thr Ser
                405                 410                 415
```

Thr Gly Asn Tyr Asn Tyr Lys Tyr Arg Tyr Leu Arg His Gly Lys Leu
420 425 430

Arg Pro Phe Glu Arg Asp Ile Ser Asn Val Pro Phe Ser Pro Asp Gly
435 440 445

Lys Pro Cys Thr Pro Pro Ala Leu Asn Cys Tyr Trp Pro Leu Asn Asp
450 455 460

Tyr Gly Phe Tyr Thr Thr Thr Gly Ile Gly Tyr Gln Pro Tyr Arg Val
465 470 475 480

Val Val Leu Ser Phe Glu Leu Leu Asn Ala Pro Ala Thr Val Cys Gly
485 490 495

Pro Lys Leu Ser Thr Asp Leu Ile Lys Asn Gln Cys Val Asn Phe Asn
500 505 510

Phe Asn Gly Leu Thr Gly Thr Gly Val Leu Thr Pro Ser Ser Lys Arg
515 520 525

Phe Gln Pro Phe Gln Gln Phe Gly Arg Asp Val Ser Asp Phe Thr Asp
530 535 540

Ser Val Arg Asp Pro Lys Thr Ser Glu Ile Leu Asp Ile Ser Pro Cys
545 550 555 560

Ala Phe Gly Gly Val Ser Val Ile Thr Pro Gly Thr Asn Ala Ser Ser
565 570 575

Glu Val Ala Val Leu Tyr Gln Asp Val Asn Cys Thr Asp Val Ser Thr
580 585 590

Ala Ile His Ala Asp Gln Leu Thr Pro Ala Trp Arg Ile Tyr Ser Thr
595 600 605

Gly Asn Asn Val Phe Gln Thr Gln Ala Gly Cys Leu Ile Gly Ala Glu
610 615 620

His Val Asp Thr Ser Tyr Glu Cys Asp Ile Pro Ile Gly Ala Gly Ile
625 630 635 640

Cys Ala Ser Tyr His Thr Val Ser Leu Leu Arg Ser Thr Ser Gln Lys
645 650 655

Ser Ile Val Ala Tyr Thr Met Ser Leu Gly Ala Asp Ser Ser Ile Ala
660 665 670

Tyr Ser Asn Asn Thr Ile Ala Ile Pro Thr Asn Phe Ser Ile Ser Ile
675 680 685

Thr Thr Glu Val Met Pro Val Ser Met Ala Lys Thr Ser Val Asp Cys
690 695 700

Asn Met Tyr Ile Cys Gly Asp Ser Thr Glu Cys Ala Asn Leu Leu Leu
705 710 715 720

Gln Tyr Gly Ser Phe Cys Thr Gln Leu Asn Arg Ala Leu Ser Gly Ile
725 730 735

Ala Ala Glu Gln Asp Arg Asn Thr Arg Glu Val Phe Ala Gln Val Lys
740 745 750

Gln Met Tyr Lys Thr Pro Thr Leu Lys Tyr Phe Gly Gly Phe Asn Phe
755 760 765
Ser Gln Ile Leu Pro Asp Pro Leu Lys Pro Thr Lys Arg Ser Phe Ile
770 775 780

<210> 48
<211> 984
<212> PRT
<213> SARS coronavirus

<400> 48
Asp Arg Cys Thr Thr Phe Asp Asp Val Gln Ala Pro Asn Tyr Thr Gln
1 5 10 15
His Thr Ser Ser Met Arg Gly Val Tyr Tyr Pro Asp Glu Ile Phe Arg
20 25 30
Ser Asp Thr Leu Tyr Leu Thr Gln Asp Leu Phe Leu Pro Phe Tyr Ser
35 40 45
Asn Val Thr Gly Phe His Thr Ile Asn His Thr Phe Gly Asn Pro Val
50 55 60
Ile Pro Phe Lys Asp Gly Ile Tyr Phe Ala Ala Thr Glu Lys Ser Asn
65 70 75 80
Val Val Arg Gly Trp Val Phe Gly Ser Thr Met Asn Asn Lys Ser Gln
85 90 95
Ser Val Ile Ile Ile Asn Asn Ser Thr Asn Val Val Ile Arg Ala Cys
100 105 110
Asn Phe Glu Leu Cys Asp Asn Pro Phe Phe Ala Val Ser Lys Pro Met
115 120 125
Gly Thr Gln Thr His Thr Met Ile Phe Asp Asn Ala Phe Asn Cys Thr
130 135 140
Phe Glu Tyr Ile Ser Asp Ala Phe Ser Leu Asp Val Ser Glu Lys Ser
145 150 155 160
Gly Asn Phe Lys His Leu Arg Glu Phe Val Phe Lys Asn Lys Asp Gly
165 170 175
Phe Leu Tyr Val Tyr Lys Gly Tyr Gln Pro Ile Asp Val Val Arg Asp
180 185 190
Leu Pro Ser Gly Phe Asn Thr Leu Lys Pro Ile Phe Lys Leu Pro Leu
195 200 205
Gly Ile Asn Ile Thr Asn Phe Arg Ala Ile Leu Thr Ala Phe Ser Pro
210 215 220
Ala Gln Asp Ile Trp Gly Thr Ser Ala Ala Ala Tyr Phe Val Gly Tyr
225 230 235 240

Leu Lys Pro Thr Thr Phe Met Leu Lys Tyr Asp Glu Asn Gly Thr Ile
245 250 255

Thr Asp Ala Val Asp Cys Ser Gln Asn Pro Leu Ala Glu Leu Lys Cys
260 265 270

Ser Val Lys Ser Phe Glu Ile Asp Lys Gly Ile Tyr Gln Thr Ser Asn
275 280 285

Phe Arg Val Val Pro Ser Gly Asp Val Val Arg Phe Pro Asn Ile Thr
290 295 300

Asn Leu Cys Pro Phe Gly Glu Val Phe Asn Ala Thr Lys Phe Pro Ser
305 310 315 320

Val Tyr Ala Trp Glu Arg Lys Lys Ile Ser Asn Cys Val Ala Asp Tyr
325 330 335

Ser Val Leu Tyr Asn Ser Thr Phe Phe Ser Thr Phe Lys Cys Tyr Gly
340 345 350

Val Ser Ala Thr Lys Leu Asn Asp Leu Cys Phe Ser Asn Val Tyr Ala
355 360 365

Asp Ser Phe Val Val Lys Gly Asp Asp Val Arg Gln Ile Ala Pro Gly
370 375 380

Gln Thr Gly Val Ile Ala Asp Tyr Asn Tyr Lys Leu Pro Asp Asp Phe
385 390 395 400

Met Gly Cys Val Leu Ala Trp Asn Thr Arg Asn Ile Asp Ala Thr Ser
405 410 415

Thr Gly Asn Tyr Asn Tyr Lys Tyr Arg Tyr Leu Arg His Gly Lys Leu
420 425 430

Arg Pro Phe Glu Arg Asp Ile Ser Asn Val Pro Phe Ser Pro Asp Gly
435 440 445

Lys Pro Cys Thr Pro Pro Ala Leu Asn Cys Tyr Trp Pro Leu Asn Asp
450 455 460

Tyr Gly Phe Tyr Thr Thr Thr Gly Ile Gly Tyr Gln Pro Tyr Arg Val
465 470 475 480

Val Val Leu Ser Phe Glu Leu Leu Asn Ala Pro Ala Thr Val Cys Gly
485 490 495

Pro Lys Leu Ser Thr Asp Leu Ile Lys Asn Gln Cys Val Asn Phe Asn
500 505 510

Phe Asn Gly Leu Thr Gly Thr Gly Val Leu Thr Pro Ser Ser Lys Arg
515 520 525

Phe Gln Pro Phe Gln Gln Phe Gly Arg Asp Val Ser Asp Phe Thr Asp
530 535 540

Ser Val Arg Asp Pro Lys Thr Ser Glu Ile Leu Asp Ile Ser Pro Cys
545 550 555 560

Ala Phe Gly Gly Val Ser Val Ile Thr Pro Gly Thr Asn Ala Ser Ser
565 570 575

Glu Val Ala Val Leu Tyr Gln Asp Val Asn Cys Thr Asp Val Ser Thr
580 585 590

Ala Ile His Ala Asp Gln Leu Thr Pro Ala Trp Arg Ile Tyr Ser Thr
595 600 605

Gly Asn Asn Val Phe Gln Thr Gln Ala Gly Cys Leu Ile Gly Ala Glu
610 615 620

His Val Asp Thr Ser Tyr Glu Cys Asp Ile Pro Ile Gly Ala Gly Ile
625 630 635 640

Cys Ala Ser Tyr His Thr Val Ser Leu Leu Arg Ser Thr Ser Gln Lys
645 650 655

Ser Ile Val Ala Tyr Thr Met Ser Leu Gly Ala Asp Ser Ser Ile Ala
660 665 670

Tyr Ser Asn Asn Thr Ile Ala Ile Pro Thr Asn Phe Ser Ile Ser Ile
675 680 685

Thr Thr Glu Val Met Pro Val Ser Met Ala Lys Thr Ser Val Asp Cys
690 695 700

Asn Met Tyr Ile Cys Gly Asp Ser Thr Glu Cys Ala Asn Leu Leu Leu
705 710 715 720

Gln Tyr Gly Ser Phe Cys Thr Gln Leu Asn Arg Ala Leu Ser Gly Ile
725 730 735

Ala Ala Glu Gln Asp Arg Asn Thr Arg Glu Val Phe Ala Gln Val Lys
740 745 750

Gln Met Tyr Lys Thr Pro Thr Leu Lys Tyr Phe Gly Gly Phe Asn Phe
755 760 765

Ser Gln Ile Leu Pro Asp Pro Leu Lys Pro Thr Lys Arg Ser Phe Ile
770 775 780

Glu Asp Leu Leu Phe Asn Lys Val Thr Leu Ala Asp Ala Gly Phe Met
785 790 795 800

Lys Gln Tyr Gly Glu Cys Leu Gly Asp Ile Asn Ala Arg Asp Leu Ile
805 810 815

Cys Ala Gln Lys Phe Asn Gly Leu Thr Val Leu Pro Pro Leu Leu Thr
820 825 830

Asp Asp Met Ile Ala Ala Tyr Thr Ala Ala Leu Val Ser Gly Thr Ala
835 840 845

Thr Ala Gly Trp Thr Phe Gly Ala Gly Ala Ala Leu Gln Ile Pro Phe
850 855 860

Ala Met Gln Met Ala Tyr Arg Phe Asn Gly Ile Gly Val Thr Gln Asn
865 870 875 880

Val Leu Tyr Glu Asn Gln Lys Gln Ile Ala Asn Gln Phe Asn Lys Ala
885 890 895

Ile Ser Gln Ile Gln Glu Ser Leu Thr Thr Thr Ser Thr Ala Leu Gly
900 905 910

Lys Leu Gln Asp Val Val Asn Gln Asn Ala Gln Ala Leu Asn Thr Leu
915 920 925
Val Lys Gln Leu Ser Ser Asn Phe Gly Ala Ile Ser Ser Val Leu Asn
930 935 940
Asp Ile Leu Ser Arg Leu Asp Lys Val Glu Ala Glu Val Gln Ile Asp
945 950 955 960
Arg Leu Ile Thr Gly Arg Leu Gln Ser Leu Gln Thr Tyr Val Thr Gln
965 970 975
Gln Leu Ile Arg Ala Ala Glu Ile
980

<210> 49
<211> 1174
<212> PRT
<213> SARS coronavirus

<400> 49
Asp Arg Cys Thr Thr Phe Asp Asp Val Gln Ala Pro Asn Tyr Thr Gln
1 5 10 15
His Thr Ser Ser Met Arg Gly Val Tyr Tyr Pro Asp Glu Ile Phe Arg
20 25 30
Ser Asp Thr Leu Tyr Leu Thr Gln Asp Leu Phe Leu Pro Phe Tyr Ser
35 40 45
Asn Val Thr Gly Phe His Thr Ile Asn His Thr Phe Gly Asn Pro Val
50 55 60
Ile Pro Phe Lys Asp Gly Ile Tyr Phe Ala Ala Thr Glu Lys Ser Asn
65 70 75 80
Val Val Arg Gly Trp Val Phe Gly Ser Thr Met Asn Asn Lys Ser Gln
85 90 95
Ser Val Ile Ile Ile Asn Asn Ser Thr Asn Val Val Ile Arg Ala Cys
100 105 110
Asn Phe Glu Leu Cys Asp Asn Pro Phe Phe Ala Val Ser Lys Pro Met
115 120 125
Gly Thr Gln Thr His Thr Met Ile Phe Asp Asn Ala Phe Asn Cys Thr
130 135 140
Phe Glu Tyr Ile Ser Asp Ala Phe Ser Leu Asp Val Ser Glu Lys Ser
145 150 155 160
Gly Asn Phe Lys His Leu Arg Glu Phe Val Phe Lys Asn Lys Asp Gly
165 170 175
Phe Leu Tyr Val Tyr Lys Gly Tyr Gln Pro Ile Asp Val Val Arg Asp
180 185 190

Leu Pro Ser Gly Phe Asn Thr Leu Lys Pro Ile Phe Lys Leu Pro Leu
195 200 205

Gly Ile Asn Ile Thr Asn Phe Arg Ala Ile Leu Thr Ala Phe Ser Pro
210 215 220

Ala Gln Asp Ile Trp Gly Thr Ser Ala Ala Ala Tyr Phe Val Gly Tyr
225 230 235 240

Leu Lys Pro Thr Thr Phe Met Leu Lys Tyr Asp Glu Asn Gly Thr Ile
245 250 255

Thr Asp Ala Val Asp Cys Ser Gln Asn Pro Leu Ala Glu Leu Lys Cys
260 265 270

Ser Val Lys Ser Phe Glu Ile Asp Lys Gly Ile Tyr Gln Thr Ser Asn
275 280 285

Phe Arg Val Val Pro Ser Gly Asp Val Val Arg Phe Pro Asn Ile Thr
290 295 300

Asn Leu Cys Pro Phe Gly Glu Val Phe Asn Ala Thr Lys Phe Pro Ser
305 310 315 320

Val Tyr Ala Trp Glu Arg Lys Lys Ile Ser Asn Cys Val Ala Asp Tyr
325 330 335

Ser Val Leu Tyr Asn Ser Thr Phe Phe Ser Thr Phe Lys Cys Tyr Gly
340 345 350

Val Ser Ala Thr Lys Leu Asn Asp Leu Cys Phe Ser Asn Val Tyr Ala
355 360 365

Asp Ser Phe Val Val Lys Gly Asp Asp Val Arg Gln Ile Ala Pro Gly
370 375 380

Gln Thr Gly Val Ile Ala Asp Tyr Asn Tyr Lys Leu Pro Asp Asp Phe
385 390 395 400

Met Gly Cys Val Leu Ala Trp Asn Thr Arg Asn Ile Asp Ala Thr Ser
405 410 415

Thr Gly Asn Tyr Asn Tyr Lys Tyr Arg Tyr Leu Arg His Gly Lys Leu
420 425 430

Arg Pro Phe Glu Arg Asp Ile Ser Asn Val Pro Phe Ser Pro Asp Gly
435 440 445

Lys Pro Cys Thr Pro Pro Ala Leu Asn Cys Tyr Trp Pro Leu Asn Asp
450 455 460

Tyr Gly Phe Tyr Thr Thr Thr Gly Ile Gly Tyr Gln Pro Tyr Arg Val
465 470 475 480

Val Val Leu Ser Phe Glu Leu Leu Asn Ala Pro Ala Thr Val Cys Gly
485 490 495

Pro Lys Leu Ser Thr Asp Leu Ile Lys Asn Gln Cys Val Asn Phe Asn
500 505 510

Phe Asn Gly Leu Thr Gly Thr Gly Val Leu Thr Pro Ser Ser Lys Arg
515 520 525

Phe Gln Pro Phe Gln Gln Phe Gly Arg Asp Val Ser Asp Phe Thr Asp
530 535 540
Ser Val Arg Asp Pro Lys Thr Ser Glu Ile Leu Asp Ile Ser Pro Cys
545 550 555 560
Ala Phe Gly Gly Val Ser Val Ile Thr Pro Gly Thr Asn Ala Ser Ser
565 570 575
Glu Val Ala Val Leu Tyr Gln Asp Val Asn Cys Thr Asp Val Ser Thr
580 585 590
Ala Ile His Ala Asp Gln Leu Thr Pro Ala Trp Arg Ile Tyr Ser Thr
595 600 605
Gly Asn Asn Val Phe Gln Thr Gln Ala Gly Cys Leu Ile Gly Ala Glu
610 615 620
His Val Asp Thr Ser Tyr Glu Cys Asp Ile Pro Ile Gly Ala Gly Ile
625 630 635 640
Cys Ala Ser Tyr His Thr Val Ser Leu Leu Arg Ser Thr Ser Gln Lys
645 650 655
Ser Ile Val Ala Tyr Thr Met Ser Leu Gly Ala Asp Ser Ser Ile Ala
660 665 670
Tyr Ser Asn Asn Thr Ile Ala Ile Pro Thr Asn Phe Ser Ile Ser Ile
675 680 685
Thr Thr Glu Val Met Pro Val Ser Met Ala Lys Thr Ser Val Asp Cys
690 695 700
Asn Met Tyr Ile Cys Gly Asp Ser Thr Glu Cys Ala Asn Leu Leu Leu
705 710 715 720
Gln Tyr Gly Ser Phe Cys Thr Gln Leu Asn Arg Ala Leu Ser Gly Ile
725 730 735
Ala Ala Glu Gln Asp Arg Asn Thr Arg Glu Val Phe Ala Gln Val Lys
740 745 750
Gln Met Tyr Lys Thr Pro Thr Leu Lys Tyr Phe Gly Gly Phe Asn Phe
755 760 765
Ser Gln Ile Leu Pro Asp Pro Leu Lys Pro Thr Lys Arg Ser Phe Ile
770 775 780
Glu Asp Leu Leu Phe Asn Lys Val Thr Leu Ala Asp Ala Gly Phe Met
785 790 795 800
Lys Gln Tyr Gly Glu Cys Leu Gly Asp Ile Asn Ala Arg Asp Leu Ile
805 810 815
Cys Ala Gln Lys Phe Asn Gly Leu Thr Val Leu Pro Pro Leu Leu Thr
820 825 830
Asp Asp Met Ile Ala Ala Tyr Thr Ala Ala Leu Val Ser Gly Thr Ala
835 840 845
Thr Ala Gly Trp Thr Phe Gly Ala Gly Ala Ala Leu Gln Ile Pro Phe
850 855 860

Ala Met Gln Met Ala Tyr Arg Phe Asn Gly Ile Gly Val Thr Gln Asn
865 870 875 880
Val Leu Tyr Glu Asn Gln Lys Gln Ile Ala Asn Gln Phe Asn Lys Ala
885 890 895
Ile Ser Gln Ile Gln Glu Ser Leu Thr Thr Thr Ser Thr Ala Leu Gly
900 905 910
Lys Leu Gln Asp Val Val Asn Gln Asn Ala Gln Ala Leu Asn Thr Leu
915 920 925
Val Lys Gln Leu Ser Ser Asn Phe Gly Ala Ile Ser Ser Val Leu Asn
930 935 940
Asp Ile Leu Ser Arg Leu Asp Lys Val Glu Ala Glu Val Gln Ile Asp
945 950 955 960
Arg Leu Ile Thr Gly Arg Leu Gln Ser Leu Gln Thr Tyr Val Thr Gln
965 970 975
Gln Leu Ile Arg Ala Ala Glu Ile Arg Ala Ser Ala Asn Leu Ala Ala
980 985 990
Thr Lys Met Ser Glu Cys Val Leu Gly Gln Ser Lys Arg Val Asp Phe
995 1000 1005
Cys Gly Lys Gly Tyr His Leu Met Ser Phe Pro Gln Ala Ala Pro His
1010 1015 1020
Gly Val Val Phe Leu His Val Thr Tyr Val Pro Ser Gln Glu Arg Asn
1025 1030 1035 1040
Phe Thr Thr Ala Pro Ala Ile Cys His Glu Gly Lys Ala Tyr Phe Pro
1045 1050 1055
Arg Glu Gly Val Phe Val Phe Asn Gly Thr Ser Trp Phe Ile Thr Gln
1060 1065 1070
Arg Asn Phe Phe Ser Pro Gln Ile Ile Thr Thr Asp Asn Thr Phe Val
1075 1080 1085
Ser Gly Asn Cys Asp Val Val Ile Gly Ile Ile Asn Asn Thr Val Tyr
1090 1095 1100
Asp Pro Leu Gln Pro Glu Leu Asp Ser Phe Lys Glu Glu Leu Asp Lys
1105 1110 1115 1120
Tyr Phe Lys Asn His Thr Ser Pro Asp Val Asp Leu Gly Asp Ile Ser
1125 1130 1135
Gly Ile Asn Ala Ser Val Val Asn Ile Gln Lys Glu Ile Asp Arg Leu
1140 1145 1150
Asn Glu Val Ala Lys Asn Leu Asn Glu Ser Leu Ile Asp Leu Gln Glu
1155 1160 1165
Leu Gly Lys Tyr Glu Gln
1170

<210> 50

<211> 260

<212> PRT

<213> SARS coronavirus

<400> 50

Asp Arg Cys Thr Thr Phe Asp Asp Val Gln Ala Pro Asn Tyr Thr Gln
1 5 10 15

His Thr Ser Ser Met Arg Gly Val Tyr Tyr Pro Asp Glu Ile Phe Arg
20 25 30

Ser Asp Thr Leu Tyr Leu Thr Gln Asp Leu Phe Leu Pro Phe Tyr Ser
35 40 45

Asn Val Thr Gly Phe His Thr Ile Asn His Thr Phe Gly Asn Pro Val
50 55 60

Ile Pro Phe Lys Asp Gly Ile Tyr Phe Ala Ala Thr Glu Lys Ser Asn
65 70 75 80

Val Val Arg Gly Trp Val Phe Gly Ser Thr Met Asn Asn Lys Ser Gln
85 90 95

Ser Val Ile Ile Ile Asn Asn Ser Thr Asn Val Val Ile Arg Ala Cys
100 105 110

Asn Phe Glu Leu Cys Asp Asn Pro Phe Phe Ala Val Ser Lys Pro Met
115 120 125

Gly Thr Gln Thr His Thr Met Ile Phe Asp Asn Ala Phe Asn Cys Thr
130 135 140

Phe Glu Tyr Ile Ser Asp Ala Phe Ser Leu Asp Val Ser Glu Lys Ser
145 150 155 160

Gly Asn Phe Lys His Leu Arg Glu Phe Val Phe Lys Asn Lys Asp Gly
165 170 175

Phe Leu Tyr Val Tyr Lys Gly Tyr Gln Pro Ile Asp Val Val Arg Asp
180 185 190

Leu Pro Ser Gly Phe Asn Thr Leu Lys Pro Ile Phe Lys Leu Pro Leu
195 200 205

Gly Ile Asn Ile Thr Asn Phe Arg Ala Ile Leu Thr Ala Phe Ser Pro
210 215 220

Ala Gln Asp Ile Trp Gly Thr Ser Ala Ala Ala Tyr Phe Val Gly Tyr
225 230 235 240

Leu Lys Pro Thr Thr Phe Met Leu Lys Tyr Asp Glu Asn Gly Thr Ile
245 250 255

Thr Asp Ala Val
260

<210> 51

<211> 430

<212> PRT

<213> SARS coronavirus

<400> 51

Asp Arg Cys Thr Thr Phe Asp Asp Val Gln Ala Pro Asn Tyr Thr Gln
1 5 10 15

His Thr Ser Ser Met Arg Gly Val Tyr Tyr Pro Asp Glu Ile Phe Arg
20 25 30

Ser Asp Thr Leu Tyr Leu Thr Gln Asp Leu Phe Leu Pro Phe Tyr Ser
35 40 45

Asn Val Thr Gly Phe His Thr Ile Asn His Thr Phe Gly Asn Pro Val
50 55 60

Ile Pro Phe Lys Asp Gly Ile Tyr Phe Ala Ala Thr Glu Lys Ser Asn
65 70 75 80

Val Val Arg Gly Trp Val Phe Gly Ser Thr Met Asn Asn Lys Ser Gln
85 90 95

Ser Val Ile Ile Ile Asn Asn Ser Thr Asn Val Val Ile Arg Ala Cys
100 105 110

Asn Phe Glu Leu Cys Asp Asn Pro Phe Phe Ala Val Ser Lys Pro Met
115 120 125

Gly Thr Gln Thr His Thr Met Ile Phe Asp Asn Ala Phe Asn Cys Thr
130 135 140

Phe Glu Tyr Ile Ser Asp Ala Phe Ser Leu Asp Val Ser Glu Lys Ser
145 150 155 160

Gly Asn Phe Lys His Leu Arg Glu Phe Val Phe Lys Asn Lys Asp Gly
165 170 175

Phe Leu Tyr Val Tyr Lys Gly Tyr Gln Pro Ile Asp Val Val Arg Asp
180 185 190

Leu Pro Ser Gly Phe Asn Thr Leu Lys Pro Ile Phe Lys Leu Pro Leu
195 200 205

Gly Ile Asn Ile Thr Asn Phe Arg Ala Ile Leu Thr Ala Phe Ser Pro
210 215 220

Ala Gln Asp Ile Trp Gly Thr Ser Ala Ala Ala Tyr Phe Val Gly Tyr
225 230 235 240

Leu Lys Pro Thr Thr Phe Met Leu Lys Tyr Asp Glu Asn Gly Thr Ile
245 250 255

Thr Asp Ala Val Asp Cys Ser Gln Asn Pro Leu Ala Glu Leu Lys Cys
260 265 270

Ser Val Lys Ser Phe Glu Ile Asp Lys Gly Ile Tyr Gln Thr Ser Asn
275 280 285

Phe Arg Val Val Pro Ser Gly Asp Val Val Arg Phe Pro Asn Ile Thr
290 295 300

Asn Leu Cys Pro Phe Gly Glu Val Phe Asn Ala Thr Lys Phe Pro Ser
305 310 315 320

Val Tyr Ala Trp Glu Arg Lys Lys Ile Ser Asn Cys Val Ala Asp Tyr
325 330 335

Ser Val Leu Tyr Asn Ser Thr Phe Phe Ser Thr Phe Lys Cys Tyr Gly
340 345 350

Val Ser Ala Thr Lys Leu Asn Asp Leu Cys Phe Ser Asn Val Tyr Ala
355 360 365

Asp Ser Phe Val Val Lys Gly Asp Asp Val Arg Gln Ile Ala Pro Gly
370 375 380

Gln Thr Gly Val Ile Ala Asp Tyr Asn Tyr Lys Leu Pro Asp Asp Phe
385 390 395 400

Met Gly Cys Val Leu Ala Trp Asn Thr Arg Asn Ile Asp Ala Thr Ser
405 410 415

Thr Gly Asn Tyr Asn Tyr Lys Tyr Arg Tyr Leu Arg His Gly
420 425 430

<210> 52
<211> 521
<212> PRT
<213> SARS coronavirus

<400> 52

Asp Arg Cys Thr Thr Phe Asp Asp Val Gln Ala Pro Asn Tyr Thr Gln
1 5 10 15

His Thr Ser Ser Met Arg Gly Val Tyr Tyr Pro Asp Glu Ile Phe Arg
20 25 30

Ser Asp Thr Leu Tyr Leu Thr Gln Asp Leu Phe Leu Pro Phe Tyr Ser
35 40 45

Asn Val Thr Gly Phe His Thr Ile Asn His Thr Phe Gly Asn Pro Val
50 55 60

Ile Pro Phe Lys Asp Gly Ile Tyr Phe Ala Ala Thr Glu Lys Ser Asn
65 70 75 80

Val Val Arg Gly Trp Val Phe Gly Ser Thr Met Asn Asn Lys Ser Gln
85 90 95

Ser Val Ile Ile Ile Asn Asn Ser Thr Asn Val Val Ile Arg Ala Cys
100 105 110

Asn Phe Glu Leu Cys Asp Asn Pro Phe Phe Ala Val Ser Lys Pro Met
115 120 125

Gly Thr Gln Thr His Thr Met Ile Phe Asp Asn Ala Phe Asn Cys Thr
130 135 140
Phe Glu Tyr Ile Ser Asp Ala Phe Ser Leu Asp Val Ser Glu Lys Ser
145 150 155 160
Gly Asn Phe Lys His Leu Arg Glu Phe Val Phe Lys Asn Lys Asp Gly
165 170 175
Phe Leu Tyr Val Tyr Lys Gly Tyr Gln Pro Ile Asp Val Val Arg Asp
180 185 190
Leu Pro Ser Gly Phe Asn Thr Leu Lys Pro Ile Phe Lys Leu Pro Leu
195 200 205
Gly Ile Asn Ile Thr Asn Phe Arg Ala Ile Leu Thr Ala Phe Ser Pro
210 215 220
Ala Gln Asp Ile Trp Gly Thr Ser Ala Ala Ala Tyr Phe Val Gly Tyr
225 230 235 240
Leu Lys Pro Thr Thr Phe Met Leu Lys Tyr Asp Glu Asn Gly Thr Ile
245 250 255
Thr Asp Ala Val Asp Cys Ser Gln Asn Pro Leu Ala Glu Leu Lys Cys
260 265 270
Ser Val Lys Ser Phe Glu Ile Asp Lys Gly Ile Tyr Gln Thr Ser Asn
275 280 285
Phe Arg Val Val Pro Ser Gly Asp Val Val Arg Phe Pro Asn Ile Thr
290 295 300
Asn Leu Cys Pro Phe Gly Glu Val Phe Asn Ala Thr Lys Phe Pro Ser
305 310 315 320
Val Tyr Ala Trp Glu Arg Lys Lys Ile Ser Asn Cys Val Ala Asp Tyr
325 330 335
Ser Val Leu Tyr Asn Ser Thr Phe Phe Ser Thr Phe Lys Cys Tyr Gly
340 345 350
Val Ser Ala Thr Lys Leu Asn Asp Leu Cys Phe Ser Asn Val Tyr Ala
355 360 365
Asp Ser Phe Val Val Lys Gly Asp Asp Val Arg Gln Ile Ala Pro Gly
370 375 380
Gln Thr Gly Val Ile Ala Asp Tyr Asn Tyr Lys Leu Pro Asp Asp Phe
385 390 395 400
Met Gly Cys Val Leu Ala Trp Asn Thr Arg Asn Ile Asp Ala Thr Ser
405 410 415
Thr Gly Asn Tyr Asn Tyr Lys Tyr Arg Tyr Leu Arg His Gly Lys Leu
420 425 430
Arg Pro Phe Glu Arg Asp Ile Ser Asn Val Pro Phe Ser Pro Asp Gly
435 440 445
Lys Pro Cys Thr Pro Pro Ala Leu Asn Cys Tyr Trp Pro Leu Asn Asp
450 455 460

Tyr Gly Phe Tyr Thr Thr Thr Gly Ile Gly Tyr Gln Pro Tyr Arg Val
465 470 475 480
Val Val Leu Ser Phe Glu Leu Leu Asn Ala Pro Ala Thr Val Cys Gly
485 490 495
Pro Lys Leu Ser Thr Asp Leu Ile Lys Asn Gln Cys Val Asn Phe Asn
500 505 510
Phe Asn Gly Leu Thr Gly Thr Gly Val
515 520

<210> 53
<211> 777
<212> PRT
<213> Artificial Sequence

<220>
<223> Synthetic sequence of amino acids 17-757 of SEQ ID NO:1 plus an N-terminal mouse K chain leader sequence and a C-terminal myc epitope and a polyhistidine tag

<400> 53
Met Glu Thr Asp Thr Leu Leu Leu Trp Val Leu Leu Leu Trp Val Pro
1 5 10 15
Gly Ser Thr Gly Asp Asp Arg Cys Thr Thr Phe Asp Asp Val Gln Ala
20 25 30
Pro Asn Tyr Thr Gln His Thr Ser Ser Met Arg Gly Val Tyr Tyr Pro
35 40 45
Asp Glu Ile Phe Arg Ser Asp Thr Leu Tyr Leu Thr Gln Asp Leu Phe
50 55 60
Leu Pro Phe Tyr Ser Asn Val Thr Gly Phe His Thr Ile Asn His Thr
65 70 75 80
Phe Gly Asn Pro Val Ile Pro Phe Lys Asp Gly Ile Tyr Phe Ala Ala
85 90 95
Thr Glu Lys Ser Asn Val Val Arg Gly Trp Val Phe Gly Ser Thr Met
100 105 110
Asn Asn Lys Ser Gln Ser Val Ile Ile Ile Asn Asn Ser Thr Asn Val
115 120 125
Val Ile Arg Ala Cys Asn Phe Glu Leu Cys Asp Asn Pro Phe Phe Ala
130 135 140
Val Ser Lys Pro Met Gly Thr Gln Thr His Thr Met Ile Phe Asp Asn
145 150 155 160
Ala Phe Asn Cys Thr Phe Glu Tyr Ile Ser Asp Ala Phe Ser Leu Asp
165 170 175

Val Ser Glu Lys Ser Gly Asn Phe Lys His Leu Arg Glu Phe Val Phe
180 185 190

Lys Asn Lys Asp Gly Phe Leu Tyr Val Tyr Lys Gly Tyr Gln Pro Ile
195 200 205

Asp Val Val Arg Asp Leu Pro Ser Gly Phe Asn Thr Leu Lys Pro Ile
210 215 220

Phe Lys Leu Pro Leu Gly Ile Asn Ile Thr Asn Phe Arg Ala Ile Leu
225 230 235 240

Thr Ala Phe Ser Pro Ala Gln Asp Ile Trp Gly Thr Ser Ala Ala Ala
245 250 255

Tyr Phe Val Gly Tyr Leu Lys Pro Thr Thr Phe Met Leu Lys Tyr Asp
260 265 270

Glu Asn Gly Thr Ile Thr Asp Ala Val Asp Cys Ser Gln Asn Pro Leu
275 280 285

Ala Glu Leu Lys Cys Ser Val Lys Ser Phe Glu Ile Asp Lys Gly Ile
290 295 300

Tyr Gln Thr Ser Asn Phe Arg Val Val Pro Ser Gly Asp Val Val Arg
305 310 315 320

Phe Pro Asn Ile Thr Asn Leu Cys Pro Phe Gly Glu Val Phe Asn Ala
325 330 335

Thr Lys Phe Pro Ser Val Tyr Ala Trp Glu Arg Lys Lys Ile Ser Asn
340 345 350

Cys Val Ala Asp Tyr Ser Val Leu Tyr Asn Ser Thr Phe Phe Ser Thr
355 360 365

Phe Lys Cys Tyr Gly Val Ser Ala Thr Lys Leu Asn Asp Leu Cys Phe
370 375 380

Ser Asn Val Tyr Ala Asp Ser Phe Val Val Lys Gly Asp Asp Val Arg
385 390 395 400

Gln Ile Ala Pro Gly Gln Thr Gly Val Ile Ala Asp Tyr Asn Tyr Lys
405 410 415

Leu Pro Asp Asp Phe Met Gly Cys Val Leu Ala Trp Asn Thr Arg Asn
420 425 430

Ile Asp Ala Thr Ser Thr Gly Asn Tyr Asn Tyr Lys Tyr Arg Tyr Leu
435 440 445

Arg His Gly Lys Leu Arg Pro Phe Glu Arg Asp Ile Ser Asn Val Pro
450 455 460

Phe Ser Pro Asp Gly Lys Pro Cys Thr Pro Pro Ala Leu Asn Cys Tyr
465 470 475 480

Trp Pro Leu Asn Asp Tyr Gly Phe Tyr Thr Thr Thr Gly Ile Gly Tyr
485 490 495

Gln Pro Tyr Arg Val Val Val Leu Ser Phe Glu Leu Leu Asn Ala Pro
500 505 510

Ala Thr Val Cys Gly Pro Lys Leu Ser Thr Asp Leu Ile Lys Asn Gln
515 520 525

Cys Val Asn Phe Asn Phe Asn Gly Leu Thr Gly Thr Gly Val Leu Thr
530 535 540

Pro Ser Ser Lys Arg Phe Gln Pro Phe Gln Gln Phe Gly Arg Asp Val
545 550 555 560

Ser Asp Phe Thr Asp Ser Val Arg Asp Pro Lys Thr Ser Glu Ile Leu
565 570 575

Asp Ile Ser Pro Cys Ala Phe Gly Gly Val Ser Val Ile Thr Pro Gly
580 585 590

Thr Asn Ala Ser Ser Glu Val Ala Val Leu Tyr Gln Asp Val Asn Cys
595 600 605

Thr Asp Val Ser Thr Ala Ile His Ala Asp Gln Leu Thr Pro Ala Trp
610 615 620

Arg Ile Tyr Ser Thr Gly Asn Asn Val Phe Gln Thr Gln Ala Gly Cys
625 630 635 640

Leu Ile Gly Ala Glu His Val Asp Thr Ser Tyr Glu Cys Asp Ile Pro
645 650 655

Ile Gly Ala Gly Ile Cys Ala Ser Tyr His Thr Val Ser Leu Leu Arg
660 665 670

Ser Thr Ser Gln Lys Ser Ile Val Ala Tyr Thr Met Ser Leu Gly Ala
675 680 685

Asp Ser Ser Ile Ala Tyr Ser Asn Asn Thr Ile Ala Ile Pro Thr Asn
690 695 700

Phe Ser Ile Ser Ile Thr Thr Glu Val Met Pro Val Ser Met Ala Lys
705 710 715 720

Thr Ser Val Asp Cys Asn Met Tyr Ile Cys Gly Asp Ser Thr Glu Cys
725 730 735

Ala Asn Leu Leu Leu Gln Tyr Gly Ser Phe Cys Thr Gln Leu Asn Arg
740 745 750

Ala Leu Ser Gly Ile Ala Ala Glu Gln Glu Gln Lys Leu Ile Ser Glu
755 760 765

Glu Asp Leu His His His His His His
770 775

<210> 54

<211> 297

<212> PRT

<213> Artificial Sequence

<220>

<223> Synthetic sequence of amino acids 17-276 of SEQ ID NO:1 plus an

N-terminal mouse K chain leader sequence and a C-terminal myc epitope and a polyhistidine tag

<400> 54

```
Met Glu Thr Asp Thr Leu Leu Leu Trp Val Leu Leu Leu Trp Val Pro
1               5                   10                  15
Gly Ser Thr Gly Asp Asp Arg Cys Thr Thr Phe Asp Asp Val Gln Ala
            20                  25                  30
Pro Asn Tyr Thr Gln His Thr Ser Ser Met Arg Gly Val Tyr Tyr Pro
        35                  40                  45
Asp Glu Ile Phe Arg Ser Asp Thr Leu Tyr Leu Thr Gln Asp Leu Phe
    50                  55                  60
Leu Pro Phe Tyr Ser Asn Val Thr Gly Phe His Thr Ile Asn His Thr
65                  70                  75                  80
Phe Gly Asn Pro Val Ile Pro Phe Lys Asp Gly Ile Tyr Phe Ala Ala
                85                  90                  95
Thr Glu Lys Ser Asn Val Val Arg Gly Trp Val Phe Gly Ser Thr Met
            100                 105                 110
Asn Asn Lys Ser Gln Ser Val Ile Ile Ile Asn Asn Ser Thr Asn Val
        115                 120                 125
Val Ile Arg Ala Cys Asn Phe Glu Leu Cys Asp Asn Pro Phe Phe Ala
    130                 135                 140
Val Ser Lys Pro Met Gly Thr Gln Thr His Thr Met Ile Phe Asp Asn
145                 150                 155                 160
Ala Phe Asn Cys Thr Phe Glu Tyr Ile Ser Asp Ala Phe Ser Leu Asp
                165                 170                 175
Val Ser Glu Lys Ser Gly Asn Phe Lys His Leu Arg Glu Phe Val Phe
            180                 185                 190
Lys Asn Lys Asp Gly Phe Leu Tyr Val Tyr Lys Gly Tyr Gln Pro Ile
        195                 200                 205
Asp Val Val Arg Asp Leu Pro Ser Gly Phe Asn Thr Leu Lys Pro Ile
    210                 215                 220
Phe Lys Leu Pro Leu Gly Ile Asn Ile Thr Asn Phe Arg Ala Ile Leu
225                 230                 235                 240
Thr Ala Phe Ser Pro Ala Gln Asp Ile Trp Gly Thr Ser Ala Ala Ala
                245                 250                 255
Tyr Phe Val Gly Tyr Leu Lys Pro Thr Thr Phe Met Leu Lys Tyr Asp
            260                 265                 270
Glu Asn Gly Thr Ile Thr Asp Ala Val Glu Gln Lys Leu Ile Ser Glu
        275                 280                 285
Glu Asp Leu His His His His His His
    290                 295
```

<210> 55

<211> 558

<212> PRT

<213> Artificial Sequence

<220>

<223> A synthetic sequence of amino acids 17-537 of SEQ ID NO:1 plus an N-terminal mouse K chain leader sequence and a C-terminal myc epitope and a polyhistidine tag

<400> 55

Met Glu Thr Asp Thr Leu Leu Leu Trp Val Leu Leu Leu Trp Val Pro
1 5 10 15

Gly Ser Thr Gly Asp Asp Arg Cys Thr Thr Phe Asp Asp Val Gln Ala
20 25 30

Pro Asn Tyr Thr Gln His Thr Ser Ser Met Arg Gly Val Tyr Tyr Pro
35 40 45

Asp Glu Ile Phe Arg Ser Asp Thr Leu Tyr Leu Thr Gln Asp Leu Phe
50 55 60

Leu Pro Phe Tyr Ser Asn Val Thr Gly Phe His Thr Ile Asn His Thr
65 70 75 80

Phe Gly Asn Pro Val Ile Pro Phe Lys Asp Gly Ile Tyr Phe Ala Ala
85 90 95

Thr Glu Lys Ser Asn Val Val Arg Gly Trp Val Phe Gly Ser Thr Met
100 105 110

Asn Asn Lys Ser Gln Ser Val Ile Ile Ile Asn Asn Ser Thr Asn Val
115 120 125

Val Ile Arg Ala Cys Asn Phe Glu Leu Cys Asp Asn Pro Phe Phe Ala
130 135 140

Val Ser Lys Pro Met Gly Thr Gln Thr His Thr Met Ile Phe Asp Asn
145 150 155 160

Ala Phe Asn Cys Thr Phe Glu Tyr Ile Ser Asp Ala Phe Ser Leu Asp
165 170 175

Val Ser Glu Lys Ser Gly Asn Phe Lys His Leu Arg Glu Phe Val Phe
180 185 190

Lys Asn Lys Asp Gly Phe Leu Tyr Val Tyr Lys Gly Tyr Gln Pro Ile
195 200 205

Asp Val Val Arg Asp Leu Pro Ser Gly Phe Asn Thr Leu Lys Pro Ile
210 215 220

Phe Lys Leu Pro Leu Gly Ile Asn Ile Thr Asn Phe Arg Ala Ile Leu
225 230 235 240

Thr Ala Phe Ser Pro Ala Gln Asp Ile Trp Gly Thr Ser Ala Ala Ala
245 250 255

Tyr Phe Val Gly Tyr Leu Lys Pro Thr Thr Phe Met Leu Lys Tyr Asp
260 265 270

Glu Asn Gly Thr Ile Thr Asp Ala Val Asp Cys Ser Gln Asn Pro Leu
275 280 285

Ala Glu Leu Lys Cys Ser Val Lys Ser Phe Glu Ile Asp Lys Gly Ile
290 295 300

Tyr Gln Thr Ser Asn Phe Arg Val Val Pro Ser Gly Asp Val Val Arg
305 310 315 320

Phe Pro Asn Ile Thr Asn Leu Cys Pro Phe Gly Glu Val Phe Asn Ala
325 330 335

Thr Lys Phe Pro Ser Val Tyr Ala Trp Glu Arg Lys Lys Ile Ser Asn
340 345 350

Cys Val Ala Asp Tyr Ser Val Leu Tyr Asn Ser Thr Phe Phe Ser Thr
355 360 365

Phe Lys Cys Tyr Gly Val Ser Ala Thr Lys Leu Asn Asp Leu Cys Phe
370 375 380

Ser Asn Val Tyr Ala Asp Ser Phe Val Val Lys Gly Asp Asp Val Arg
385 390 395 400

Gln Ile Ala Pro Gly Gln Thr Gly Val Ile Ala Asp Tyr Asn Tyr Lys
405 410 415

Leu Pro Asp Asp Phe Met Gly Cys Val Leu Ala Trp Asn Thr Arg Asn
420 425 430

Ile Asp Ala Thr Ser Thr Gly Asn Tyr Asn Tyr Lys Tyr Arg Tyr Leu
435 440 445

Arg His Gly Lys Leu Arg Pro Phe Glu Arg Asp Ile Ser Asn Val Pro
450 455 460

Phe Ser Pro Asp Gly Lys Pro Cys Thr Pro Pro Ala Leu Asn Cys Tyr
465 470 475 480

Trp Pro Leu Asn Asp Tyr Gly Phe Tyr Thr Thr Thr Gly Ile Gly Tyr
485 490 495

Gln Pro Tyr Arg Val Val Val Leu Ser Phe Glu Leu Leu Asn Ala Pro
500 505 510

Ala Thr Val Cys Gly Pro Lys Leu Ser Thr Asp Leu Ile Lys Asn Gln
515 520 525

Cys Val Asn Phe Asn Phe Asn Gly Leu Thr Gly Thr Gly Val Glu Gln
530 535 540

Lys Leu Ile Ser Glu Glu Asp Leu His His His His His His
545 550 555

<210> 56
<211> 739
<212> PRT
<213> Artificial Sequence

<220>
<223> A synthetic sequence of amino acids 17-756 of SEQ ID NO:1 without a signal peptide at the N-terminus

<400> 56

Asp Arg Cys Thr Thr Phe Asp Asp Val Gln Ala Pro Asn Tyr Thr Gln
1 5 10 15

His Thr Ser Ser Met Arg Gly Val Tyr Tyr Pro Asp Glu Ile Phe Arg
20 25 30

Ser Asp Thr Leu Tyr Leu Thr Gln Asp Leu Phe Leu Pro Phe Tyr Ser
35 40 45

Asn Val Thr Gly Phe His Thr Ile Asn His Thr Phe Gly Asn Pro Val
50 55 60

Ile Pro Phe Lys Asp Gly Ile Tyr Phe Ala Ala Thr Glu Lys Ser Asn
65 70 75 80

Val Val Arg Gly Trp Val Phe Gly Ser Thr Met Asn Asn Lys Ser Gln
85 90 95

Ser Val Ile Ile Ile Asn Asn Ser Thr Asn Val Val Ile Arg Ala Cys
100 105 110

Asn Phe Glu Leu Cys Asp Asn Pro Phe Phe Ala Val Ser Lys Pro Met
115 120 125

Gly Thr Gln Thr His Thr Met Ile Phe Asp Asn Ala Phe Asn Cys Thr
130 135 140

Phe Glu Tyr Ile Ser Asp Ala Phe Ser Leu Asp Val Ser Glu Lys Ser
145 150 155 160

Gly Asn Phe Lys His Leu Arg Glu Phe Val Phe Lys Asn Lys Asp Gly
165 170 175

Phe Leu Tyr Val Tyr Lys Gly Tyr Gln Pro Ile Asp Val Val Arg Asp
180 185 190

Leu Pro Ser Gly Phe Asn Thr Leu Lys Pro Ile Phe Lys Leu Pro Leu
195 200 205

Gly Ile Asn Ile Thr Asn Phe Arg Ala Ile Leu Thr Ala Phe Ser Pro
210 215 220

Ala Gln Asp Ile Trp Gly Thr Ser Ala Ala Ala Tyr Phe Val Gly Tyr
225 230 235 240

Leu Lys Pro Thr Thr Phe Met Leu Lys Tyr Asp Glu Asn Gly Thr Ile
245 250 255

Thr Asp Ala Val Asp Cys Ser Gln Asn Pro Leu Ala Glu Leu Lys Cys
260 265 270

Ser Val Lys Ser Phe Glu Ile Asp Lys Gly Ile Tyr Gln Thr Ser Asn
275 280 285

Phe Arg Val Val Pro Ser Gly Asp Val Val Arg Phe Pro Asn Ile Thr
290 295 300

Asn Leu Cys Pro Phe Gly Glu Val Phe Asn Ala Thr Lys Phe Pro Ser
305 310 315 320

Val Tyr Ala Trp Glu Arg Lys Lys Ile Ser Asn Cys Val Ala Asp Tyr
325 330 335

Ser Val Leu Tyr Asn Ser Thr Phe Phe Ser Thr Phe Lys Cys Tyr Gly
340 345 350

Val Ser Ala Thr Lys Leu Asn Asp Leu Cys Phe Ser Asn Val Tyr Ala
355 360 365

Asp Ser Phe Val Val Lys Gly Asp Asp Val Arg Gln Ile Ala Pro Gly
370 375 380

Gln Thr Gly Val Ile Ala Asp Tyr Asn Tyr Lys Leu Pro Asp Asp Phe
385 390 395 400

Met Gly Cys Val Leu Ala Trp Asn Thr Arg Asn Ile Asp Ala Thr Ser
405 410 415

Thr Gly Asn Tyr Asn Tyr Lys Tyr Arg Tyr Leu Arg His Gly Lys Leu
420 425 430

Arg Pro Phe Glu Arg Asp Ile Ser Asn Val Pro Phe Ser Pro Asp Gly
435 440 445

Lys Pro Cys Thr Pro Pro Ala Leu Asn Cys Tyr Trp Pro Leu Asn Asp
450 455 460

Tyr Gly Phe Tyr Thr Thr Thr Gly Ile Gly Tyr Gln Pro Tyr Arg Val
465 470 475 480

Val Val Leu Ser Phe Glu Leu Leu Asn Ala Pro Ala Thr Val Cys Gly
485 490 495

Pro Lys Leu Ser Thr Asp Leu Ile Lys Asn Gln Cys Val Asn Phe Asn
500 505 510

Phe Asn Gly Leu Thr Gly Thr Gly Val Leu Thr Pro Ser Ser Lys Arg
515 520 525

Phe Gln Pro Phe Gln Gln Phe Gly Arg Asp Val Ser Asp Phe Thr Asp
530 535 540

Ser Val Arg Asp Pro Lys Thr Ser Glu Ile Leu Asp Ile Ser Pro Cys
545 550 555 560

Ala Phe Gly Gly Val Ser Val Ile Thr Pro Gly Thr Asn Ala Ser Ser
565 570 575

Glu Val Ala Val Leu Tyr Gln Asp Val Asn Cys Thr Asp Val Ser Thr
580 585 590

Ala Ile His Ala Asp Gln Leu Thr Pro Ala Trp Arg Ile Tyr Ser Thr
595 600 605
Gly Asn Asn Val Phe Gln Thr Gln Ala Gly Cys Leu Ile Gly Ala Glu
610 615 620
His Val Asp Thr Ser Tyr Glu Cys Asp Ile Pro Ile Gly Ala Gly Ile
625 630 635 640
Cys Ala Ser Tyr His Thr Val Ser Leu Leu Arg Ser Thr Ser Gln Lys
645 650 655
Ser Ile Val Ala Tyr Thr Met Ser Leu Gly Ala Asp Ser Ser Ile Ala
660 665 670
Tyr Ser Asn Asn Thr Ile Ala Ile Pro Thr Asn Phe Ser Ile Ser Ile
675 680 685
Thr Thr Glu Val Met Pro Val Ser Met Ala Lys Thr Ser Val Asp Cys
690 695 700
Asn Met Tyr Ile Cys Gly Asp Ser Thr Glu Cys Ala Asn Leu Leu Leu
705 710 715 720
Gln Tyr Gly Ser Phe Cys Thr Gln Leu Asn Arg Ala Leu Ser Gly Ile
725 730 735
Ala Ala Glu

<210> 57
<211> 265
<212> PRT
<213> Artificial Sequence

<220>
<223> A synthetic sequence of amino acids 272-537 of SEQ ID NO:1

<400> 57
Ile Thr Asp Ala Val Asp Cys Ser Gln Asn Pro Leu Ala Glu Leu Lys
1 5 10 15
Cys Ser Val Lys Ser Phe Glu Ile Asp Lys Gly Ile Tyr Gln Thr Ser
20 25 30
Asn Phe Arg Val Val Pro Ser Gly Asp Val Val Arg Phe Asn Ile Thr
35 40 45
Asn Leu Cys Pro Phe Gly Glu Val Phe Asn Ala Thr Lys Phe Pro Ser
50 55 60
Val Tyr Ala Trp Glu Arg Lys Lys Ile Ser Asn Cys Val Ala Asp Tyr
65 70 75 80
Ser Val Leu Tyr Asn Ser Thr Phe Phe Ser Thr Phe Lys Cys Tyr Gly
85 90 95

Val Ser Ala Thr Lys Leu Asn Asp Leu Cys Phe Ser Asn Val Tyr Ala
100 105 110
Asp Ser Phe Val Val Lys Gly Asp Asp Val Arg Gln Ile Ala Pro Gly
115 120 125
Gln Thr Gly Val Ile Ala Asp Tyr Asn Tyr Lys Leu Pro Asp Asp Phe
130 135 140
Met Gly Cys Val Leu Ala Trp Asn Thr Arg Asn Ile Asp Ala Thr Ser
145 150 155 160
Thr Gly Asn Tyr Asn Tyr Lys Tyr Arg Tyr Leu Arg His Gly Lys Leu
165 170 175
Arg Pro Phe Glu Arg Asp Ile Ser Asn Val Pro Phe Ser Pro Asp Gly
180 185 190
Lys Pro Cys Thr Pro Pro Ala Leu Asn Cys Tyr Trp Pro Leu Asn Asp
195 200 205
Tyr Gly Phe Tyr Thr Thr Thr Gly Ile Gly Tyr Gln Pro Tyr Arg Val
210 215 220
Val Val Leu Ser Phe Glu Leu Leu Asn Ala Pro Ala Thr Val Cys Gly
225 230 235 240
Pro Lys Leu Ser Thr Asp Leu Ile Lys Asn Gln Cys Val Asn Phe Asn
245 250 255
Phe Asn Gly Leu Thr Gly Thr Gly Val
260 265

<210> 58
<211> 17
<212> PRT
<213> SARS coronavirus

<400> 58
Asp Val Gln Ala Pro Asn Tyr Thr Gln His Thr Ser Ser Met Arg Gly
1 5 10 15
Cys

<210> 59
<211> 15
<212> PRT
<213> SARS coronavirus

<400> 59
Pro Ser Ser Lys Arg Phe Gln Pro Gln Gln Phe Gly Arg Asp Cys
1 5 10 15

<210> 60
<211> 16
<212> PRT
<213> SARS coronavirus

<400> 60
Met Phe Ile Phe Leu Leu Phe Leu Thr Leu Thr Ser Gly Ser Asp Leu
1 5 10 15

<210> 61
<211> 235
<212> PRT
<213> Artificial Sequence

<220>
<223> A synthetic sequence of amino acids 303-537 of SEQ ID NO:1
containing the receptor binding domain

<400> 61
Ser Asn Phe Arg Val Val Pro Ser Gly Asp Val Val Arg Phe Pro Asn
1 5 10 15
Ile Thr Asn Leu Cys Pro Phe Gly Glu Val Phe Asn Ala Thr Lys Phe
20 25 30
Pro Ser Val Tyr Ala Trp Glu Arg Lys Lys Ile Ser Asn Cys Val Ala
35 40 45
Asp Tyr Ser Val Leu Tyr Asn Ser Thr Phe Phe Ser Thr Phe Lys Cys
50 55 60
Tyr Gly Val Ser Ala Thr Lys Leu Asn Asp Leu Cys Phe Ser Asn Val
65 70 75 80
Tyr Ala Asp Ser Phe Val Val Lys Gly Asp Asp Val Arg Gln Ile Ala
85 90 95
Pro Gly Gln Thr Gly Val Ile Ala Asp Tyr Asn Tyr Lys Leu Pro Asp
100 105 110
Asp Phe Met Gly Cys Val Leu Ala Trp Asn Thr Arg Asn Ile Asp Ala
115 120 125
Thr Ser Thr Gly Asn Tyr Asn Tyr Lys Tyr Arg Tyr Leu Arg His Gly
130 135 140
Lys Leu Arg Pro Phe Glu Arg Asp Ile Ser Asn Val Pro Phe Ser Pro
145 150 155 160
Asp Gly Lys Pro Cys Thr Pro Pro Ala Leu Asn Cys Tyr Trp Pro Leu
165 170 175

Asn Asp Tyr Gly Phe Tyr Thr Thr Thr Gly Ile Gly Tyr Gln Pro Tyr
180 185 190

Arg Val Val Val Leu Ser Phe Glu Leu Leu Asn Ala Pro Ala Thr Val
195 200 205

Cys Gly Pro Lys Leu Ser Thr Asp Leu Ile Lys Asn Gln Cys Val Asn
210 215 220

Phe Asn Phe Asn Gly Leu Thr Gly Thr Gly Val
225 230 235

<210> 62
<211> 199
<212> PRT
<213> Artificial Sequence

<220>
<223> A synthetic sequence of amino acids 319-517 of SEQ ID NO:1 containing the receptor binding domain

<400> 62

Ile Thr Asn Leu Cys Pro Phe Gly Glu Val Phe Asn Ala Thr Lys Phe
1 5 10 15

Pro Ser Val Tyr Ala Trp Glu Arg Lys Lys Ile Ser Asn Cys Val Ala
20 25 30

Asp Tyr Ser Val Leu Tyr Asn Ser Thr Phe Phe Ser Thr Phe Lys Cys
35 40 45

Tyr Gly Val Ser Ala Thr Lys Leu Asn Asp Leu Cys Phe Ser Asn Val
50 55 60

Tyr Ala Asp Ser Phe Val Val Lys Gly Asp Asp Val Arg Gln Ile Ala
65 70 75 80

Pro Gly Gln Thr Gly Val Ile Ala Asp Tyr Asn Tyr Lys Leu Pro Asp
85 90 95

Asp Phe Met Gly Cys Val Leu Ala Trp Asn Thr Arg Asn Ile Asp Ala
100 105 110

Thr Ser Thr Gly Asn Tyr Asn Tyr Lys Tyr Arg Tyr Leu Arg His Gly
115 120 125

Lys Leu Arg Pro Phe Glu Arg Asp Ile Ser Asn Val Pro Phe Ser Pro
130 135 140

Asp Gly Lys Pro Cys Thr Pro Pro Ala Leu Asn Cys Tyr Trp Pro Leu
145 150 155 160

Asn Asp Tyr Gly Phe Tyr Thr Thr Thr Gly Ile Gly Tyr Gln Pro Tyr
165 170 175

Arg Val Val Val Leu Ser Phe Glu Leu Leu Asn Ala Pro Ala Thr Val
180 185 190

Cys Gly Pro Lys Leu Ser Thr
195

<210> 63

<211> 200

<212> PRT

<213> Artificial Sequence

<220>

<223> A synthetic sequence of amino acids 319-518 of SEQ ID NO:1 containing the receptor binding domain

<400> 63

Ile Thr Asn Leu Cys Pro Phe Gly Glu Val Phe Asn Ala Thr Lys Phe
1 5 10 15

Pro Ser Val Tyr Ala Trp Glu Arg Lys Lys Ile Ser Asn Cys Val Ala
20 25 30

Asp Tyr Ser Val Leu Tyr Asn Ser Thr Phe Phe Ser Thr Phe Lys Cys
35 40 45

Tyr Gly Val Ser Ala Thr Lys Leu Asn Asp Leu Cys Phe Ser Asn Val
50 55 60

Tyr Ala Asp Ser Phe Val Val Lys Gly Asp Asp Val Arg Gln Ile Ala
65 70 75 80

Pro Gly Gln Thr Gly Val Ile Ala Asp Tyr Asn Tyr Lys Leu Pro Asp
85 90 95

Asp Phe Met Gly Cys Val Leu Ala Trp Asn Thr Arg Asn Ile Asp Ala
100 105 110

Thr Ser Thr Gly Asn Tyr Asn Tyr Lys Tyr Arg Tyr Leu Arg His Gly
115 120 125

Lys Leu Arg Pro Phe Glu Arg Asp Ile Ser Asn Val Pro Phe Ser Pro
130 135 140

Asp Gly Lys Pro Cys Thr Pro Pro Ala Leu Asn Cys Tyr Trp Pro Leu
145 150 155 160

Asn Asp Tyr Gly Phe Tyr Thr Thr Thr Gly Ile Gly Tyr Gln Pro Tyr
165 170 175

Arg Val Val Val Leu Ser Phe Glu Leu Leu Asn Ala Pro Ala Thr Val
180 185 190

Cys Gly Pro Lys Leu Ser Thr Asp
195 200

<210> 64
<211> 23
<212> DNA
<213> Artificial Sequence

<220>
<223> A synthetic primer

<400> 64
gatcggatcc ggtacaatca cag 23

<210> 65
<211> 23
<212> DNA
<213> Artificial Sequence

<220>
<223> A synthetic primer

<400> 65
gatcgggccc gacacactgg ttc 23

INTERNATIONAL SEARCH REPORT

International Application No
PCT/US2004/023345

A. CLASSIFICATION OF SUBJECT MATTER
IPC 7 C07K14/165 A61K39/215 A61K39/42 A61K38/16

According to International Patent Classification (IPC) or to both national classification and IPC

B. FIELDS SEARCHED

Minimum documentation searched (classification system followed by classification symbols)
IPC 7 C07K A61K

Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched

Electronic data base consulted during the international search (name of data base and, where practical, search terms used)

EPO-Internal, BIOSIS, WPI Data, EMBASE, MEDLINE, PAJ, Sequence Search

C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category *	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
X	DATABASE EMBL 23 April 2003 (2003-04-23), "SARS coronavirus urbani, complete genome." XP002304795 retrieved from EBI Database accession no. AY278741 abstract	1-84
X	-& ROTA P A ET AL: "Characterization of a novel coronavirus associated with severe acute respiratory syndrome" SCIENCE, AMERICAN ASSOCIATION FOR THE ADVANCEMENT OF SCIENCE,, US, vol. 300, no. 5624, 30 May 2003 (2003-05-30), pages 1394-1399, XP002269482 ISSN: 0036-8075 the whole document	1-84

-/--

[X] Further documents are listed in the continuation of box C.

[X] Patent family members are listed in annex.

° Special categories of cited documents :

"A" document defining the general state of the art which is not considered to be of particular relevance

"E" earlier document but published on or after the international filing date

"L" document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified)

"O" document referring to an oral disclosure, use, exhibition or other means

"P" document published prior to the international filing date but later than the priority date claimed

"T" later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention

"X" document of particular relevance; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step when the document is taken alone

"Y" document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art

"&" document member of the same patent family

Date of the actual completion of the international search	Date of mailing of the international search report
10 November 2004	26/11/2004

Name and mailing address of the ISA	Authorized officer
European Patent Office, P.B. 5818 Patentlaan 2 NL - 2280 HV Rijswijk Tel. (+31-70) 340-2040, Tx. 31 651 epo nl, Fax: (+31-70) 340-3016	Grötzinger, T

1

Form PCT/ISA/210 (second sheet) (January 2004)

C.(Continuation) DOCUMENTS CONSIDERED TO BE RELEVANT

Category *	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
X	DATABASE EMBL 15 April 2003 (2003-04-15), "SARS coronavirus TOR2, complete genome." XP002304796 retrieved from EBI Database accession no. AY274119 abstract	1-84
X	-& MARRA M A ET AL: "The genome sequence of the SARS-associated coronavirus" SCIENCE, AMERICAN ASSOCIATION FOR THE ADVANCEMENT OF SCIENCE,, US, vol. 300, no. 5624, 30 May 2003 (2003-05-30), pages 1399-1404, XP002269483 ISSN: 0036-8075 the whole document	1-84
X	WO 93/23421 A (SMITHKLINE BEECHAM CORP ; JONES ELAINE V (US); KLEPFER SHARON (US); MI) 25 November 1993 (1993-11-25) page 3, line 24 - line 31 page 4, line 6 - line 30	39,40, 44-48, 59-61, 67-69
P,X	DATABASE EMBL 30 November 2003 (2003-11-30), "SARS coronavirus CUHK-AG01, complete genome." XP002304797 retrieved from EBI Database accession no. AY345986 abstract	1-84
P,X	-& CHIM S ET AL: "Genomic characterisation of the severe acute respiratory syndrome coronavirus of Amoy Gardens outbreak in Hong Kong" LANCET THE, LANCET LIMITED. LONDON, GB, vol. 362, no. 9398, 29 November 2003 (2003-11-29), pages 1807-1808, XP004476558 ISSN: 0140-6736 the whole document	1-84
P,X	DATABASE EMBL 7 January 2004 (2004-01-07), "SARS coronavirus TW6, complete genome." XP002304798 retrieved from EBI Database accession no. AY502929 abstract	1-84

-/--

1

C.(Continuation) DOCUMENTS CONSIDERED TO BE RELEVANT

Category *	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
P,X	-& YEH SHIOU-HWEI ET AL: "Characterization of severe acute respiratory syndrome coronavirus genomes in Taiwan: Molecular epidemiology and genome evolution." PROCEEDINGS OF THE NATIONAL ACADEMY OF SCIENCES OF THE UNITED STATES OF AMERICA, vol. 101, no. 8, 24 February 2004 (2004-02-24), pages 2542-2547, XP002304793 ISSN: 0027-8424 the whole document	1-84
P,X	DATABASE EMBL 28 February 2004 (2004-02-28), "SARS coronavirus TW-GD5 isolate TW-GD5_SC22-23 replicase 1B and spike glycoprotein genes, partial cds." XP002304799 retrieved from EBI Database accession no. AY451903 abstract	1,4-20, 24-30, 37-48, 67-69, 73,74
P,X	YANG ZHI-YONG ET AL: "A DNA vaccine induces SARS coronavirus neutralization and protective immunity in mice" NATURE (LONDON), vol. 428, no. 6982, 1 April 2004 (2004-04-01), pages 561-564, XP002304794 ISSN: 0028-0836 abstract page 563, right-hand column, section "Immunogen and plasmid construction"	1-84

1

INTERNATIONAL SEARCH REPORT

International application No.
PCT/US2004/023345

Box No. I Nucleotide and/or amino acid sequence(s) (Continuation of item 1.b of the first sheet)

1. With regard to any nucleotide and/or amino acid sequence disclosed in the international application and necessary to the claimed invention, the international search was carried out on the basis of:

   a. type of material
      - [X] a sequence listing
      - [ ] table(s) related to the sequence listing

   b. format of material
      - [X] in written format
      - [X] in computer readable form

   c. time of filing/furnishing
      - [X] contained in the international application as filed
      - [X] filed together with the international application in computer readable form
      - [ ] furnished subsequently to this Authority for the purpose of search

2. [ ] In addition, in the case that more than one version or copy of a sequence listing and/or table relating thereto has been filed or furnished, the required statements that the information in the subsequent or additional copies is identical to that in the application as filed or does not go beyond the application as filed, as appropriate, were furnished.

3. Additional comments:

INTERNATIONAL SEARCH REPORT

International application No.
PCT/US2004/023345

Box II Observations where certain claims were found unsearchable (Continuation of item 2 of first sheet)

This International Search Report has not been established in respect of certain claims under Article 17(2)(a) for the following reasons:

1. [X] Claims Nos.:
because they relate to subject matter not required to be searched by this Authority, namely:

Although claims 59, 62, 66, 67, as well as the dependent claims are directed to a method of treatment of the human/animal body, the search has been carried out and based on the alleged effects of the compound/composition.

2. [ ] Claims Nos.:
because they relate to parts of the International Application that do not comply with the prescribed requirements to such an extent that no meaningful International Search can be carried out, specifically:

3. [ ] Claims Nos.:
because they are dependent claims and are not drafted in accordance with the second and third sentences of Rule 6.4(a).

Box III Observations where unity of invention is lacking (Continuation of item 3 of first sheet)

This International Searching Authority found multiple inventions in this international application, as follows:

1. [ ] As all required additional search fees were timely paid by the applicant, this International Search Report covers all searchable claims.

2. [ ] As all searchable claims could be searched without effort justifying an additional fee, this Authority did not invite payment of any additional fee.

3. [ ] As only some of the required additional search fees were timely paid by the applicant, this International Search Report covers only those claims for which fees were paid, specifically claims Nos.:

4. [ ] No required additional search fees were timely paid by the applicant. Consequently, this International Search Report is restricted to the invention first mentioned in the claims; it is covered by claims Nos.:

Remark on Protest

[ ] The additional search fees were accompanied by the applicant's protest.

[ ] No protest accompanied the payment of additional search fees.

INTERNATIONAL SEARCH REPORT

Information on patent family members

International Application No
PCT/US2004/023345

Patent document cited in search report	Publication date	Patent family member(s)	Publication date
WO 9323421 A	25-11-1993	AU 678970 B2	19-06-1997
		AU 4240493 A	13-12-1993
		AU 678971 B2	19-06-1997
		AU 4241093 A	13-12-1993
		CA 2134898 A1	25-11-1993
		CA 2135201 A1	25-11-1993
		EP 0640096 A1	01-03-1995
		EP 0640097 A1	01-03-1995
		JP 7508176 T	14-09-1995
		JP 8501931 T	05-03-1996
		WO 9323421 A1	25-11-1993
		WO 9323422 A1	25-11-1993